

Resolving the Flexibility and Intricacy of DNA Repair Protein-DNA Interactions

Andrew Martin Cobb

May 2010

Thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy.

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that no quotation from the thesis, nor any information derived therefrom, may be published without the author's prior, written consent.

Acknowledgements

Firstly, I would like to thank my primary supervisor Dr Richard Bowater for his encouragement and support throughout this study.

I am also very grateful for the guidance and advice that I received from my secondary supervisors Dr Julea Butt and Dr Helen James during the numerous progress meetings that were held.

I would also like to acknowledge Dr Tracey Swingler, Dr David Widdick, Dr Brian Jackson, Dr Heather Sayer, Dr Andrew Hemmings, Eileen Gallagher and Verity Lyall for their technical assistance.

I would like to say thanks to my parents Stephen and Jenny, my sister Sarah and also Melis for always being there for me.

Lastly, I am grateful to the BBSRC for funding the work presented here.

Abstract

Within all cells, complex molecular systems exist that are responsible for maintaining genome stability by detecting and repairing dangerous alterations in DNA. Ensuring the accurate and efficient functioning of such systems is necessary for the preservation of DNA integrity and avoidance of disease.

The flexible and diverse modes of DNA-binding exhibited by human p53 permits this 'guardian of the genome' to elicit versatile cellular activities that are crucial in monitoring threats to genome dynamics and conducting appropriate responses. In conjunction with its sequence-specific DNA-binding activity that is essential to target gene transactivation, p53 can bind to unusual DNA structures independent of DNA sequence and it has been proposed this activity may allow p53 to interact with detrimental secondary structures that arise in unstable genomic regions. To provide further insight into p53-DNA interactions, an *in vitro* DNA binding assay was developed that was used to characterise binding properties towards several DNA molecules to allow comparison of non-specific, sequence-specific and structure-specific binding. It was determined that unusual structures in DNA significantly enhanced p53 binding in non-sequence specific DNA and that the presence of internal hairpin regions induced binding comparable to sequence-specific binding. *In vivo* p53-DNA interactions were also quantified using chromatin immunoprecipitation and variations in preference to different response element sequences was ascertained.

DNA binding is also central to the ability of Ku proteins to function as essential components of non-homologous end joining and telomere maintenance in eukaryotes. Prokaryotic homologues of Ku proteins that function as homodimers in two-component repair systems have also been identified. Recently, 3 Ku homologues in *Streptomyces coelicolor* were reported, but very little is currently known regarding their biological activity. It was discovered that all 3 Ku proteins exhibited varied independent DNA-binding properties that were influenced by DNA topology, size and end-structure. Unusually for Ku, it was found 1 of these proteins exhibited strong binding to single-stranded DNA. Precipitation assays determined that these proteins may act as DNA end synapsis mediators during the DNA end-joining process and ligation experiments revealed Ku was responsible for rigidifying DNAs or completely inhibiting ligation activity, probably via DNA end-protection activity. Experimental evidence indicated that specific interactions could occur between *S. coelicolor* Ku suggesting these proteins form both homodimers and heterodimers.

Table of Contents

Acknowledgements	i
Abstract	ii
Table of Contents	iii
Abbreviations	vii
CHAPTER 1 - Introduction	1
1.1 Structure and Function of DNA and Proteins.	1
1.2 Protein-DNA Interactions.	3
1.3 DNA Damage and Repair.	7
1.3.1 DNA Damage.	7
1.3.2 Non-B DNA.	8
1.3.3 Triplet Repeat Sequences.	11
1.3.4 DNA Repair Pathways.	14
1.3.5 Double-strand Break Repair.	16
1.4 Protein Components of Non-homologous End Joining.	18
1.4.1 Activities of Ku Proteins.	18
1.4.2 Current Understanding of Prokaryotic NHEJ and Ku.	20
1.4.3 Ku proteins of <i>Streptomyces coelicolor</i> .	24
1.4.4 Non-homologous End Joining DNA Ligase Proteins.	27
1.5 p53 Tumour Suppressor Protein.	28
1.5.1 p53 Structure and Function.	28
1.5.2 Downstream Events following p53 Activation.	32
1.5.3 p53-DNA Interactions.	34
1.5.4 p53 and Non-B DNA.	36
1.6 Aims of this Study.	38
CHAPTER 2 – Materials and Methods	40
2.1 Growth and Maintenance of Bacterial Strains.	40
2.1.1 Bacterial Strains.	40
2.1.2 Growth of Bacterial Cultures.	41
2.2 DNA Procedures.	41
2.2.1 Transformation of Plasmid DNA into <i>E. coli</i> .	41
2.2.2 Bacterial Plasmids and Cosmids used During Studies.	42
2.2.3 Plasmid Extraction from <i>E. coli</i> .	43
2.2.4 Clean-up of Extracted Plasmids.	43
2.2.5 Polymerase Chain Reaction (PCR).	43
2.2.5.1 Amplification of Promoter Regions Following Chromatin Immunoprecipitation.	43

2.2.5.2 Amplification of DNA for Cloning.	44
2.2.5.3 Colony PCR.	45
2.2.6 Agarose Gel Electrophoresis.	45
2.2.7 DNA Purification Following Agarose Gel Electrophoresis.	45
2.2.8 Cloning of <i>scku</i> Genes.	46
2.2.8.1 Cloning <i>scku</i> Genes into pCR [®] -Blunt II-TOPO [®] .	46
2.2.8.2 DNA Sequencing.	46
2.2.8.3 Cloning <i>scku</i> Genes into Integration Vectors.	46
2.3 Oligonucleotide Preparation and Analysis.	47
2.3.1 Preparing DNA Substrates for <i>in vitro</i> DNA Binding Assays.	47
2.3.2 Analysis of Oligonucleotides by Native and Denaturing Gel Electrophoresis.	49
2.3.3 Transfer to Membrane and Chemiluminescent Detection of Oligonucleotides.	50
2.4 Protein Over-expression and Purification.	50
2.4.1 pET Expression System.	50
2.4.2 Protein Over-expression.	51
2.4.3 Protein Purification.	51
2.4.3.1 Protein Extraction.	51
2.4.3.2 Novagen His·Bind [®] Column Purification.	51
2.4.3.3 GE Healthcare HisTrap [™] HP Purification.	52
2.4.3.4 Desalting and Buffer Exchange.	52
2.4.3.5 Fast Performance Liquid Chromatography (FPLC).	53
2.4.3.6 Homogenisation of High Five Insect Cells.	53
2.5 Bacterial Cell Extract Analysis.	53
2.5.1 <i>E. coli</i> and <i>S. coelicolor</i> Cell Extract Harvesting.	53
2.5.2 <i>S. coelicolor</i> Spore Harvesting, Viability Count and Extract Analysis.	54
2.6 Protein Analysis.	54
2.6.1 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE).	54
2.6.2 Coomassie Staining of SDS-PAGE Gels.	55
2.6.3 Western Blotting.	55
2.6.4 Protein Concentration Determination using Bradford Assay.	56
2.6.5 Calculating Protein Absorbance at 280 nm.	56
2.7 <i>In vitro</i> Streptavidin Magnetic Bead (SMB) DNA Binding Assay.	56
2.7.1 <i>In vitro</i> SMB DNA Binding Assay Principle and Mechanism.	56
2.7.2 SMB Capacity Analysis.	58
2.7.3 Dissociation Constant (K _d) Equation used During SMB DNA Binding Assays.	58
2.8 <i>In vitro</i> Protein-G Bead (PGB) Assay.	59
2.9 <i>In vitro</i> Co-Immunoprecipitation (Co-IP) Assay.	60
2.10 <i>In vitro</i> Ligation Assay.	61
2.11 Analytical Ultracentrifugation.	61
2.12 Dynamic Light Scattering.	61
2.13 Chromatin Immunoprecipitation (ChIP).	62
2.13.1 Cross-linked Chromatin Preparation.	62
2.13.2 ChIP assay.	62
2.14 Antibody Production for ScKu1 and ScKu2.	63
2.15 Media and Buffers.	63
2.15.1 Media.	63

2.15.2 Additions to Bacterial Culture Media.	64
2.15.3 Buffers Used in DNA Procedures.	64
2.15.4 Buffers Used during Protein Procedures.	64
2.15.5 Buffers used during <i>in vitro</i> Assays.	65
2.15.6 Buffers Used during <i>in vivo</i> Assays.	65
CHAPTER 3 – Characterising Human p53-DNA Interactions	66
3.1 Human p53 Protein Analysis.	68
3.1.1 Insect Cell Derived Human p53 Extraction and Analysis.	68
3.1.2 Human Recombinant p53-B Analysis.	69
3.2 Oligonucleotide Annealing and Analysis.	71
3.2.1 Optimisation of Annealing Procedure.	72
3.2.2 Production of Unusual Double-stranded Oligonucleotides.	74
3.3 Dynabeads® M-280 Streptavidin and Dynabeads® Protein-G Analysis.	76
3.3.1 Dynabeads® M-280 Streptavidin Analysis.	76
3.3.1.1 Capacity Analysis.	76
3.3.1.2 Non-specific Protein-bead Interactions.	79
3.3.2 Dynabeads® Protein-G Analysis.	80
3.4 <i>In vitro</i> SMB DNA Binding Analysis of Human p53.	82
3.4.1 How Ionic Strength Affects Human p53-DNA Interactions.	84
3.4.2 Comparing Human p53s Derived From Eukaryotic and Prokaryotic Host cells.	86
3.4.3 Human p53-Interactions with Unusual DNAs.	89
3.5 Chromatin Immunoprecipitation.	92
3.5.1 Chromatin Analysis.	93
3.5.2 <i>In vivo</i> p53-DNA Interactions.	94
3.6 Discussion.	97
CHAPTER 4 – Verification of Prokaryotic Ku and Analysis of DNA Binding	103
4.1 Purification and Analysis of Prokaryotic Ku Proteins.	104
4.1.1 Novagen His-Bind® Column Purification.	105
4.1.2 GE Healthcare HisTrap™ HP Purification.	105
4.1.3 Coomassie and Western Blot Analysis.	107
4.2 MALDI-TOF Peptide Mass Fingerprinting.	109
4.3 <i>In vitro</i> SMB DNA Binding Analysis of Prokaryotic NHEJ Components.	110
4.3.1 Investigating Ku and DNA Ligase Binding to dsDNA and ssDNA.	111
4.3.2 ScKu Competition Assays.	114
4.4 <i>In vitro</i> Protein-G Bead (PGB) Assay.	118
4.4.1 Exploring the Rapid Degradation Properties of ScKu Proteins.	119
4.4.2 ScKu-Plasmid and Oligonucleotide DNA Interactions.	120
4.4.3 The Temperature Sensitive DNA Binding Activity of MtKu.	124
4.4.4 Prokaryotic Ku Interactions with Relaxed Circular DNA.	126
4.5 The Prokaryotic Ku Influence on <i>in vitro</i> Ligation.	127
4.5.1 T4 DNA Ligase Ligation.	127
4.5.2 Determining How Prokaryotic Ku Influence Ligation by T4 DNA Ligase.	129
4.6 Discussion.	131

CHAPTER 5 – Macromolecular Interactions Involving ScKu Proteins	141
5.1 Cloning of Recombinant <i>sku1</i> and <i>sku2</i> into Integration Phage Vectors.	142
5.1.1 Restriction Digest Analysis of Integration Vectors pMS82 and pSET152.	143
5.1.2 PCR Amplification of Recombinant <i>sku1</i> and <i>sku2</i> .	144
5.1.3 Zero Blunt® TOPO® Cloning of Recombinant <i>sku2</i> .	146
5.1.4 Cloning of <i>sku2</i> into pMS82 and pSET152.	149
5.2 ScKu1 and ScKu2 Immunisation and Antibody Analysis.	150
5.2.1 Analysis of Recombinant ScKu1 and ScKu2 using Polyclonal Antibodies.	150
5.2.2 <i>S. coelicolor</i> and <i>E. coli</i> Cell Extract Analysis using ScKu and MtrA Antibodies.	153
5.2.3 <i>S. coelicolor</i> Spore Extract Analysis using ScKu Polyclonal Antibodies.	156
5.3 Resolving ScKu Protein Macromolecular Interactions.	157
5.3.1 Dynamic Light Scattering Analysis of ScKu.	158
5.3.2 Co-Immunoprecipitation of <i>S. coelicolor</i> NHEJ Components.	159
5.3.3 <i>In vitro</i> SMB DNA and Protein Interaction Analysis.	161
5.4 Discussion.	165
 CHAPTER 6 – General Discussion	 170
6.1 Characterising Human p53-DNA Interactions.	170
6.2 Future Human p53-DNA Interaction Investigations.	176
6.3 Determining Prokaryotic Ku Function.	177
6.4 Future ScKu NHEJ Investigations.	182
 References	 185
 Appendix I	 197
 Appendix II	 198
 Appendix III	 199
 Appendix IV	 199
 Appendix V	 200

Abbreviations

A	Adenine
AMP	Adenosine mono-phosphate
ADP	Adenosine di-phosphate
ATP	Adenosine tri-phosphate
AUC	Analytical ultracentrifuge
BER	Base excision repair
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
C	Cytosine
CDBD	Central DNA binding domain
CDK	Cyclin-dependent kinase
ChIP	Chromatin immunoprecipitation
Co-IP	Co-immunoprecipitation
CTD	Carboxyl terminal domain
DM	Myotonic dystrophy
DNA	Deoxyribonucleic acid
DNA-PKcs	DNA protein kinase catalytic subunit
dNTP	Deoxyribonucleotide tri-phosphate
DSB	Double strand break
dsDNA	Double-stranded DNA
EMSA	Electrophoretic mobility shift assay
FHB	Final harvest bleed
FPLC	Fast performance liquid chromatography
G	Guanine
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GG-NER	Global genome nucleotide excision repair
HEH	Helix-extended region-helix
HPV	Human papilloma virus
hMSH2	MutS homologue 2
HP	Hairpin
HR	Homologous recombination
IPTG	Isopropyl- β -D-thiogalactopyranoside
LB	Luria-Bertani

Abbreviations Continued

MALDI	Matrix-assisted laser desorption/ionisation
Mdm2	Mouse double mutant 2
MMR	Mismatch repair
NER	Nucleotide excision repair
NHEJ	Nonhomologous end-joining
NTD	Amino-terminal domain
OD	Optical density
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PGB	Protein-G bead
RE	Recognition element
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SMB	Streptavidin magnetic bead
ssDNA	Single-stranded DNA
T	Thymine
TC-NER	Transcription coupled nucleotide excision repair
TFIIH	Transcription factor II-H
TOF	Time-of-flight
TRS	Trinucleotide repeat sequence
UEA	University of East Anglia
UV	Ultraviolet
WC	Watson-Crick

CHAPTER 1

Introduction

1.1 Structure and Function of DNA and Proteins.

Genetic information encoded by deoxyribonucleic acid (DNA) is used in the development and orchestration of cellular functions of most living organisms. A well recognised activity of DNA is as a long-term storage molecule that contains the instructions necessary for the production of other cellular components including proteins and RNA molecules. DNA is generally composed of 2 anti-parallel polymers of nucleotides that are supported by a phosphodiester backbone. The nucleotides consist of sugar molecules and the nitrogenous bases adenine, thymine, guanine and cytosine, of which coalesce into various combinations to produce precise genetic codes that are read in a 5' to 3' directional manner. The structural properties of DNA dictate its function and hence govern extensive cellular characteristics. In most instances, DNA in biological systems adopts the renowned anti-parallel right-handed double-helix, referred to as B-DNA (Watson and Crick, 1953), that results in a charged structure with a hydrophobic centre and hydrophilic backbone stabilised by hydrogen bonds (Bates and Maxwell, 2005). B-DNA exhibits both major and minor grooves that are 22 Å and 12 Å wide respectively. Factors such as DNA sequence, chemical modification, hydration and supercoiling can induce structural alterations so that DNA adopts non-B conformations including A-DNA (alternate groove sizes) and Z-DNA (left-handed double-helix that is implicated in transcription regulation (Oh et al., 2002)). Other unusual conformations of DNA include quadruplex formation, branched DNA, hairpin structures, Holliday junctions and single-stranded DNA (ssDNA) that are described in more detail later in this chapter.

As stated, a key role of DNA is the storage of genetic information that can be used to produce protein molecules following transcription and translation events. These biological macromolecules are essential to all organisms and participate in a myriad of cellular processes. Proteins are organic compounds that consist of various amino acid residues dictated by the gene sequence of bases of the transcribed DNA. Amino acid residues are commonly subject to chemical post-translational modifications that ultimately influence protein folding, stability and activity. The large majority of proteins naturally fold into native three-dimensional conformations that are crucial for activity, a process that can occur

unassisted or via the cooperation of molecular chaperones (Kroll, 2002). Overall protein structure is generally dictated by 4 principle aspects that are illustrated in Figure 1.1.

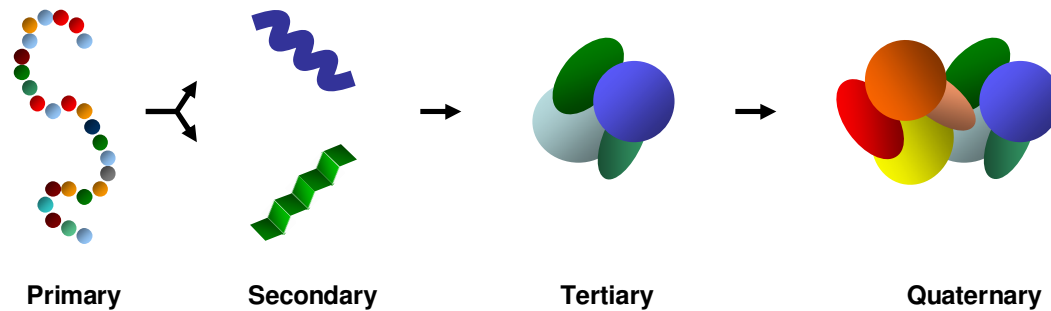


Figure 1.1 The 4 levels of complexity that determine protein structure and function. The primary determinant of a protein is the sequence of amino acid residues that are arranged in a linear chain connected by peptide bonds. The secondary structure of a protein is formation of local structures such as alpha helices and beta sheets that chains of amino acids can adopt. The tertiary structure of a protein is the overall three-dimensional shape of the protein that is determined by the spatial relationship between secondary structures and is responsible for dictating its basic function. Quaternary structure is the formation of a multi-polypeptide complex that functions as a single protein complex.

The sequence of amino acid residues and any post-translational modifications determine the formation of local secondary structures that are stabilised by hydrogen bonds. One such structure is the alpha helix whereby the protein polypeptide backbone is tightly wound around a central axis with amino acid R groups protruding outwards. Polypeptides can also extend into zigzag structures known as beta sheets that as with alpha helices are dependent on the amino acid composition of the protein. Other important secondary structures include beta turns that comprise connecting elements between alpha helices and beta sheets. The spatial arrangement of these local structures is referred to as a proteins tertiary structure that is stabilised by several elements including formation of a hydrophobic core, salt bridges, hydrogen bonds and disulfide bonds (Nelson and Cox, 2000). Amino acids that are relatively far apart in protein sequence may directly interact in the fully folded three-dimensional protein conformation. The fourth and final level of protein structure is the quaternary structure whereby multiple polypeptide chains or subunits combine into three-dimensional multicomplexes. Simple protein multimers can be composed of 2 identical (homodimers) or different (heterodimers) proteins whereas some functional proteins require much larger complexes that contain hundreds of protein subunits.

An important characteristic of several proteins is the ability to directly interact with DNA and to modulate various cellular processes including DNA replication, transcription and DNA repair and maintenance. Protein-DNA interactions show high variation in specificity and

flexibility with certain proteins, such as structural proteins, able to bind non-specifically to any DNA sequence, and other proteins, such as transcription factors, only able to bind to precise genomic regions. Other DNA-binding proteins can detect and interact with specific types of DNA damage that can be a significant threat to cell or organism viability. As each of the $\sim 10^{13}$ cells in the human body receives tens of thousands of insults to DNA every day (Jackson and Bartek, 2009), the efficient detection and removal processes performed by these specialised proteins is essential for maintaining genomic integrity.

1.2 Protein-DNA Interactions.

As our current knowledge regarding protein-nucleic acid complex assembly continues to expand, the diversity and intricacy involved in these interactions becomes increasingly apparent. Prior to formation of macromolecular protein-DNA complexes, the individual formation of specific and homogenous three-dimensional structures of both protein and DNA components are required. The thermodynamic tendency to bury nonpolar residues into the interior of a protein is a principal factor in stabilising protein folding. Similarly, restricting the surface exposure of planar surfaces of nucleic acid bases by base stacking is instrumental in duplex DNA stabilisation (Calladine, 1982). The resulting conformation exposes polar and charged residues on the DNA surface and buries nonpolar groups internally (von Hippel, 2007). The energies required for the formation of these folded entities must exceed the opposing force that is the decrease in configurational entropy. Crucially, the specificity of protein folding means individual types of amino acid residues are likely to be located at precise positions within the folded macromolecule (Hartl and Hayer-Hartl, 2009). Internal hydrogen-bonding alignments provide fundamental stability to the tertiary protein structure in conjunction with dipole-dipole and van der Waals interactions.

As mentioned, proteins can further assemble into multi-subunit complexes consisting of other proteins and also with nucleic acid components through interactions with specific intermolecular interfaces. Electrostatic forces often significantly influences these interactions alongside hydrophobic bonding and, in some instances, ligand or ion binding (Teilum et al., 2009). The equilibrium stability of multi-subunit complexes is also dependent on the concentration of the individual components and the solvent environment. Additionally, solvents, such as ethanol, can be added to the aqueous solution to promote protein-protein and protein-nucleic acid interactions and salts also have significant effects on the solvent conditions (von Hippel, 2007). Crowding by other non-interacting molecular components such as synthetic polymers or BSA can help to stabilise multi-subunit complexes by occupying elements of the solution that assemblies could potentially unfold or dissociate into

(Batra et al., 2009). Intracellular salt concentrations vary between organisms but generally fluctuate around 1 – 200 mM (Dominy et al., 2002). The highly negatively charged phosphate backbone of DNA destabilises nucleic acid structures and promotes denaturing of double-stranded DNA (dsDNA). Addition of salt neutralises this negative repulsive charge and so nucleic acid structures are stabilised at optimal ionic strengths (Schildkraut, 1965). The effect of salt on protein structure can either be stabilising or destabilising, depending on the specific charge distribution within the protein (Kohn et al., 1997).

Protein-DNA interactions are generally flexible in nature and this is evident as upon protein binding, the DNA structure is often significantly perturbed yet the protein can still remain bound (N'Soukpoe-Kossi et al., 2008). Supercoiling, local unwinding and base-pair breathing of DNA can occur following attachment of protein partners and this must also be endured by the protein in order to maintain the complex. Perhaps the best understood mode of protein-DNA interaction is the binding of a protein to a specific nucleic acid sequence. Proteins that bind in this sequence-specific manner are often able to overcome slight base-pair alterations in the DNA (Coulocheri et al., 2007), again demonstrating considerable flexibility.

Specific binding can be defined as a molecular association in which a particular molecule is bound tightly and exclusively in an energetically and kinetically stabilised complex (Oda and Nakamura, 2000). This type of interaction is crucial for the functions of various regulatory proteins that act at precise locations on the genome. During sequence-specific DNA binding, protein interactions with specific nucleic acid sequences is primarily determined by recognition of hydrogen-bonding determinants situated in the major and minor grooves of the dsDNA that interact with complementary recognition amino acids of the protein itself (Seeman et al., 1976). Hence, interactions are promoted when DNA bases are arranged in an optimal sequence. The hydrogen bond donor and acceptor patterns in the dsDNA grooves are recognised by complementary hydrogen bond donor and acceptors on the protein surface. The protein displaces water molecules and forms interfaces of complementary hydrogen-bonding patterns that are separated from the surrounding aqueous environment (von Hippel, 1994; von Hippel, 2007). Electrostatic interactions between the hydrophilic and negatively charged DNA backbone and positively charged and dipolar amino acids provide energy that stabilises the recognition surfaces. This type of interaction is heavily determined by structure and flexibility of both the DNA and protein. Although sequence-specific proteins associate tightly with recognition sites, the strength of binding does not prevent the protein from eventually dissociating from the substrate DNA once the downstream events of the interaction have been processed (Rhodes et al., 1996).

As described, certain proteins have strong binding affinity for target sequences, however most, if not all, DNA binding proteins have the ability to bind non-specifically to DNA. This binding can be said to be a random molecular association as there is no exclusive determinant required. The main component providing the necessary free energy to stabilise the interaction is the electrostatic attraction between positively charged amino acids and the negatively charged phosphodiester DNA backbone (Jen-Jacobson, 1997). This electrostatic, non-specific binding affinity, based on the displacement of counter-ions from the DNA, is not as tight as the previously described sequence-specific interaction. Consequently, proteins are able to move along DNA (von Hippel, 1994) by thermal motion in an exploratory fashion until the target binding site is located. Location of unique sites in the genome by regulatory proteins that are present in limited numbers would be considerably slower without the ability to translocate along DNA via this non-specific interaction. The process of protein diffusion in reduced dimensions includes short-range hopping along the same DNA strand and direct interstrand transfer as well as one-dimensional sliding (Figure 1.2). These processes reduce the volume through which the protein has to conduct its search for its target (Kabata et al., 1993; Kampmann, 2004; Winter et al., 1981). Upon recognition of the target sequence, proteins undergo reversible conformational changes and change from a non-specific complex dominated by electrostatic interactions to a conformation specific for tight association with target DNA base pairs (Coulocheri et al., 2007). A general consensus is that the rate of target site location is optimal when protein diffusion involves both sliding and hoping movements in combination (Halford, 2009).

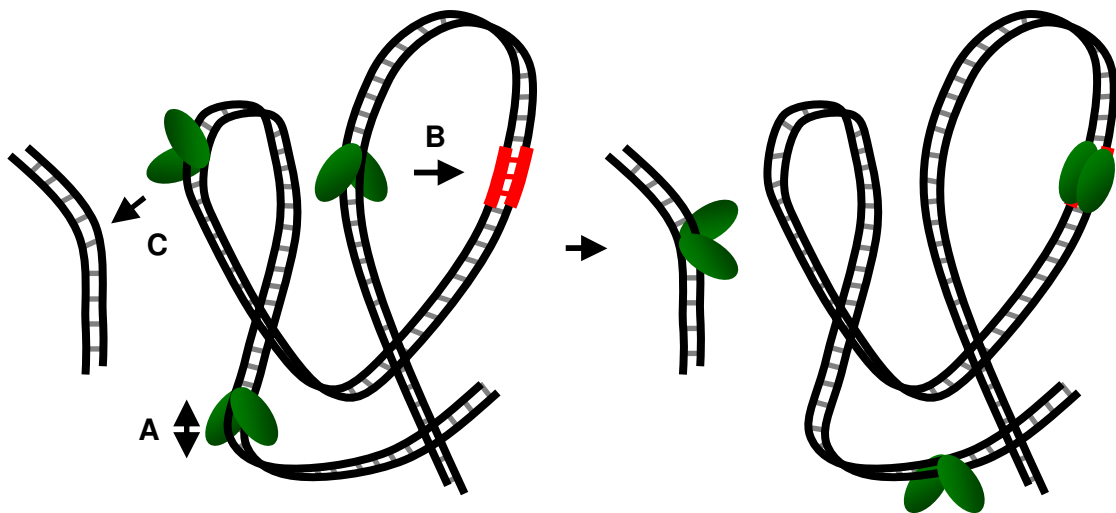


Figure 1.2 Diffusion in reduced dimensions. DNA binding proteins can interact non-specifically with DNA, which facilitates location of specific target sites (in red) by reducing excursions of the protein away from the DNA. Three modes of protein translocation along nonspecific DNA are shown: **A.** One-dimensional sliding. **B.** Intrastrand hopping. **C.** Interstrand transfer. The protein undergoes a conformational change once bound to specific DNA resulting in a tighter association.

It is recognised that the structure of DNA directly influences sequence-specific protein binding as molecular interfaces between interacting proteins and DNA are complementary in shape, allowing a close-fit association of the protein surface to the structure of the DNA (Pabo and Nekludova, 2000; Yeh et al., 2003). This complementation in shape is determined by chemical contacts including hydrogen bonding, electrostatic interactions, van der Waals forces and hydrophobic interactions. DNA structure is not only a significant determinant of sequence-specific protein-DNA complexes, but certain proteins can recognise and bind with strong affinity to specific DNA structures in a sequence-independent fashion. Such DNA regions can be essential specialised sites that require a non-B structure, some arise accidentally during various cellular processes and others can be damage induced and can be very detrimental.

An example of an important structure-dependent protein-DNA interaction involves the recognition of single-stranded genomic regions that can be important for DNA metabolism. Denaturing of B-form DNA can occur following thermal fluctuations that induce base-pair breathing whereby dsDNA opens and closes spontaneously (McConnell and von Hippel, 1970). Specialised ssDNA binding proteins interact specifically with open single-stranded regions, independent of sequence, and act to stabilise the ssDNA and allow initiation of events such as replication (Kur et al., 2005). Similarly to nonspecific binding, the predominant component of this interaction is electrostatic.

Certain damage inducing reagents are capable of creating double-strand breaks (DSBs) in DNA that can significantly reduce genomic stability. As discussed below, DNA repair mechanisms have developed that are capable of recognising and repairing resulting DNA ends. However, in order to prevent catastrophic chromosomal rearrangements, it is essential that the natural ends of linear chromosomes must be excluded from this activity (Riha et al., 2006). This is achieved by the formation of unique high order nucleoprotein structures called telomeres that consist of extended repeat DNA sequences bound by specific proteins. In most eukaryotes, the G-rich repeat forms a 3' single-stranded overhang that mediates formation of a complex structure, referred to as a t-loop. Although many of the proteins that interact with telomeres are believed to be sequence-dependent, the telomere-end protection family of proteins bind to the 3' single-stranded region independent of sequence (Croy and Wuttke, 2006).

Various structurally and functionally unrelated proteins have been shown to interact tightly with four-way Holliday junctions in DNA that arise during recombination and replication-related events. Resolving of these junctions is essential to allow DNA to conform back into 2

separate duplexes. This separation of Holliday junctions is performed by resolvase proteins such as GEN1 (Ip et al., 2008) that specifically recognise and bind to four-way junctions, again sequence-independently. Aside from proteins involved in recombination, HMG1 box proteins, SWI/SNF complex and unrelated prokaryotic proteins also recognise and bind to four-way junctions (Lilley, 1997). The ability to interact with bent DNA and capacity to increase DNA bending following binding (Zlatanova and van Holde, 1998) are common properties of such proteins.

Although recent studies have enhanced our understanding of the wide scope of protein-nucleic acid interactions, several important areas remain poorly understood. The energetic nature of protein-DNA interactions during conformational changes, the dynamic recognition of sequence-specific DNA by proteins and the interactions between protein-DNA complexes with water and counter-ions are yet to be fully understood.

1.3 DNA Damage and Repair.

1.3.1 DNA Damage.

Cells are continuously subjected to potentially catastrophic insults throughout their lifetimes that can lead to mutations, disease and death. The primary structure of DNA is continuously modified by endogenous and exogenous DNA damaging agents that can cause various alterations, ranging from base damage and nucleotide substitutions to more complex chromosomal rearrangements. The majority of DNA damage affects the primary structure of DNA, usually via chemical modifications to nucleotides, which can introduce errors including non-native base chemical bonding and formation of bulky adducts that perturb the helical structure (Kulkarni and Wilson, 2008). Exogenous damaging agents are physical or chemical agents in the external environment, including ultraviolet (UV) light, ionising radiation and extremes in pH and temperature. Sources of endogenous agents derive from the cell itself and include alkylating metabolites, water, reactive oxygen species and reactive nitrogen species (Lombard et al., 2005). DNA can also become damaged when errors occur during cellular processes that result in DNA alterations. In eukaryotes, and to an extent in prokaryotes, damage to DNA can be recognised by global surveillance mechanisms that initiate biochemical pathways such as DNA repair, cell cycle arrest, alterations in gene expression and induction of apoptosis in multi-cellular organisms (Figure 1.3).

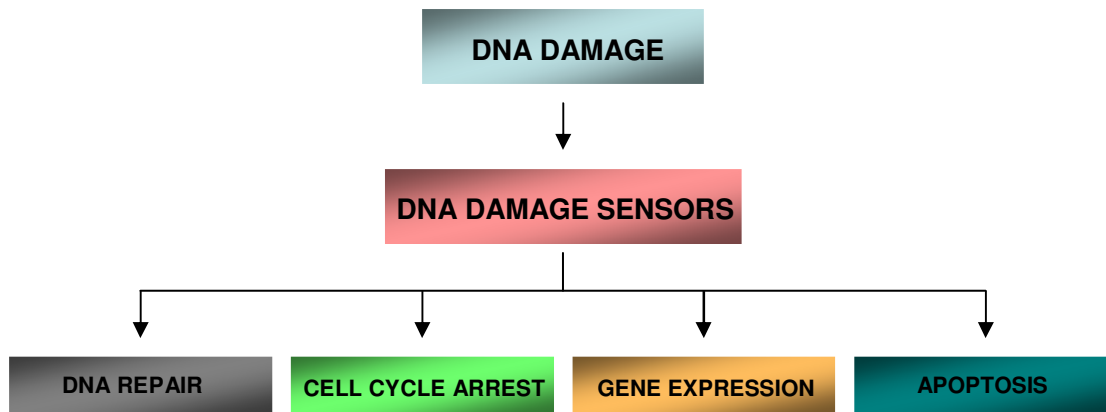


Figure 1.3 Responses to DNA damage in mammalian cells. These responses may function independently but often extensive cross-talk occurs between processes, usually by proteins participating in more than 1 response.

Defects in any of the responses shown in Figure 1.3 can compromise genomic stability. Recognition of DNA damage is the fundamental step in the initiation of subsequent downstream responses and this is often performed directly by surveillance proteins that have complementarity towards particular DNA damage induced structures (Sancar et al., 2004).

1.3.2 Non-B DNA.

DNA itself can be a causative factor for genomic instability. The three-dimensional conformations adopted by polymorphic sections of DNA can lead to the expansion of repeat tracts or the induction of gross genomic rearrangements that can cause severe diseases in humans (Bacolla and Wells, 2009). Although the majority of DNA exists in an orthodox B-form as a right-handed anti-parallel double helix with Watson-Crick (WC) base pairing, some sequences of DNA can exist in multiple conformations (Figure 1.4). It is predicted that more than ten non-B conformations can be formed at specific sequence motifs as a function of negative supercoiling density (Bacolla and Wells, 2004). Structures such as hairpins, cruciforms and Z-DNA maintain WC base pairing, however other non-B DNA molecules, including tetraplexes and slipped-strand structures, contain non-WC base pairs. These unusual structures often have higher energy states that are maintained by free energy derived from negative supercoiling, alongside other factors (Bacolla and Wells, 2009).

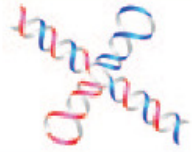
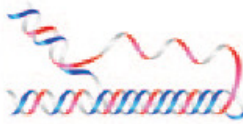

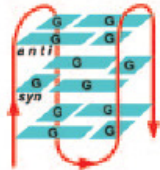

Name	Conformation	General Seq. Requirements	Sequence
Cruciform		Inverted Repeats	$\begin{array}{c} \text{TCGGTACCGA} \\ \text{AGCCATGGCT} \end{array}$
Triplex		$(R \cdot Y)_n$ Mirror Repeats	$\begin{array}{c} \text{AAGAGGGGAGAA} \\ \text{TTCTCCCTCTT} \end{array}$
Slipped (Hairpin) Structure		Direct Repeats	$\begin{array}{c} \text{TCGGTTCGGT} \\ \text{AGCCAAGCCA} \end{array}$
Tetraplex		Oligo (G) _n Tracts	$\text{AG}_3(\text{T}_2\text{AG}_3)_3$ single strand
Left-handed Z-DNA		$(YR \cdot YR)_n$	$\begin{array}{c} \text{CGCGTGCGTGTG} \\ \text{GCGCACGCACAC} \end{array}$

Figure 1.4 Non-B DNA structures involved in genetic instability. Natural genomic sequences can conform to unusual structures that demonstrate characteristics that differ from standard B-DNA. These unusual structures can significantly reduce genomic integrity and can also induce chromosomal rearrangements. Taken from (Bacolla et al., 2004).

Various non-B DNA structures arise at repeating sequences that naturally occur in genomes that can vary substantially in size, location and sequence composition (Wang and Vasquez, 2009). Hairpins and cruciforms form within inverted repeat sequences, when the duplex denatures to ssDNA during events such as replication, transcription or repair. These events provide an opportunity for a single-stranded inverted repeat sequence to form intramolecular base pairs with itself, forming a new B-helix capped by a single-stranded loop. Cruciform structures are highly comparable to Holliday junction recombination intermediates, with the four-way junction either in an open conformation to allow migration, or in a locked stacked assembly (Bacolla and Wells, 2004; Watson et al., 2004). These types of structures can be very detrimental to processes acting on DNA and can induce damage including DSBs and translocations (Collins et al., 1982).

Polypurine-polypyrimidine repeating elements have the propensity to adopt unusual triplex DNA structures consisting of 3 DNA strands. This occurs when part of the duplex denatures

and allows 1 of the single-strands to pair with the purine strand in the remaining duplex via Hoogsteen hydrogen bonds in the major groove (Wang and Vasquez, 2009). Sites containing triplex DNA show a significant increase in mutation frequency (Wang and Vasquez, 2004).

Nucleic acid sequences rich in guanines, such as telomeric DNA, can adopt a tetrad structure, consisting of square arrays of 4 guanine residues. Each guanine donates 2 Hoogsteen hydrogen bonds to an adjacent base, and receives 2 bonds from another in a cyclic arrangement (Bacolla and Wells, 2004). Four-stranded DNA is composed of multiple stacked guanine tetrads and so requires sequences that contain at least 4 copies of multiple guanines. The resulting structures are highly polymorphic and can vary regarding polarity, groove size, and strand stoichiometry. Alternative base-pairing has also been observed in some *in vitro* tetrads (Zhang et al., 2001).

Left-handed Z-DNA forms in sequences containing alternating pyrimidine-purine base-pairs. The bases flip 180° through a rotation of purines from *anti* to *syn* conformation. Z-DNA only contains 1 type of groove, unlike B-DNA that contains both major and minor grooves (Watson et al., 2004).

Tandem repeat sequences are prone to strand-slippage, where looped-out bases form duplexes stabilised by unusual base-pairing. These structures can cause contraction or expansion of tandem repeats, events that are associated with at least 20 human neurological diseases (Myers, 2004; Rass et al., 2007; Wells et al., 2005; Wojciechowska et al., 2005). Repeat tract expansion is also related to fragile sites in the genome that are predisposed to DNA breaks, chromosomal translocations and amplifications. Repeat sequence expansion and consequences are discussed further in Section 1.3.3.

Breakpoints of gross chromosomal rearrangements including translocations and deletions coincide with sequences prone to adopt non-B DNA structures, demonstrating that they can act as initiators for such potentially catastrophic events. Slipped-strand structures, triplexes, tetraplexes and Z-DNA are able to elicit dramatic genetic alterations by recombination and repair mechanisms (Bacolla and Wells, 2009). The resulting rearrangements of genetic material are the foundation for many human genomic diseases.

1.3.3 Triplet Repeat Sequences.

A direct genetic link exists between defects in DNA repair pathways and some neurological abnormalities. Defects in repair pathways can lead to loss of genomic integrity, which can consequently result in severe outcomes, including apoptosis. In non-replicating fully differentiated cells, such as neuronal tissue, apoptosis induced by genomic instability can lead to tissue degeneration and atrophy (Kulkarni and Wilson, 2008; Rass et al., 2007). Repetitive sequences are common within eukaryotic genomes and are regions that are often subjected to recombination and random integrations. Therefore, these regions are thought to have been instrumental during genome evolution for the creation of sequence diversity, however, some alterations in these regions can significantly perturb homeostasis of genome dynamics (Kovtun and McMurray, 2008). Trinucleotide repeat sequences (TRS) are of considerable importance as genomic amplifications of TRS cause various human disorders, including neuromuscular and neurodegenerative diseases (Gatchel and Zoghbi, 2005).

Hereditary neurological diseases in humans that result from TRS expansion include Huntington's disease, myotonic dystrophy, fragile X syndrome and several spinocerebellar ataxias (Bowater and Wells, 2001). For all TRS diseases, the disease phenotype severity increases in proportion to the TRS expansion length. Most triplet repeat diseases also demonstrate anticipation, whereby the disease is more severe and has an earlier age of onset with successive family generations, which is attributed to incorporation of longer repeats during transmission to the offspring. TRS diseases have been characterised into 2 groups; Type 1 diseases, such as Huntington's disease, are determined by relatively small TRS expansions (less than 100) located in coding regions of genes and large polyamino acid tracts (usually glutamine) are produced. Type 2 diseases, including myotonic dystrophy (DM), result from expansions that are significantly larger than those of type 1, and they usually occur in untranslated regions of mRNA (Lee and Kim, 2006). Expansions and deletions are thought to arise in TRS following errors during DNA metabolism that result in slippage of either the complementary strand or the newly synthesised strand (Pearson et al., 2002). Initial studies suggested unrepaired strand slippage in the nascent DNA strand could be converted into expanded repeats following a second round of DNA replication. However, it was not known how this happened, why not all repeat tracts expanded and why slippage usually only caused limited repeat length polymorphisms (Mirkin, 2006).

Once TRS tracts exceed a threshold of around 35 repeats, the tract becomes highly prone to further expansions (Lee et al., 1999) that can conform to non-B DNA that may adversely affect genomic stability (Figure 1.5). These secondary structures can be incorporated during

DNA processing events, but the exact process of this secondary structure formation remains poorly understood. Denaturing and renaturing of TRS, such as during DNA synthesis in replication, transcription or repair, is accepted to be important to create unusually stable hairpin structures that kinetically trap repetitive DNA in this otherwise unfavourable conformation (Panigrahi et al., 2005; Pearson and Sinden, 1996).

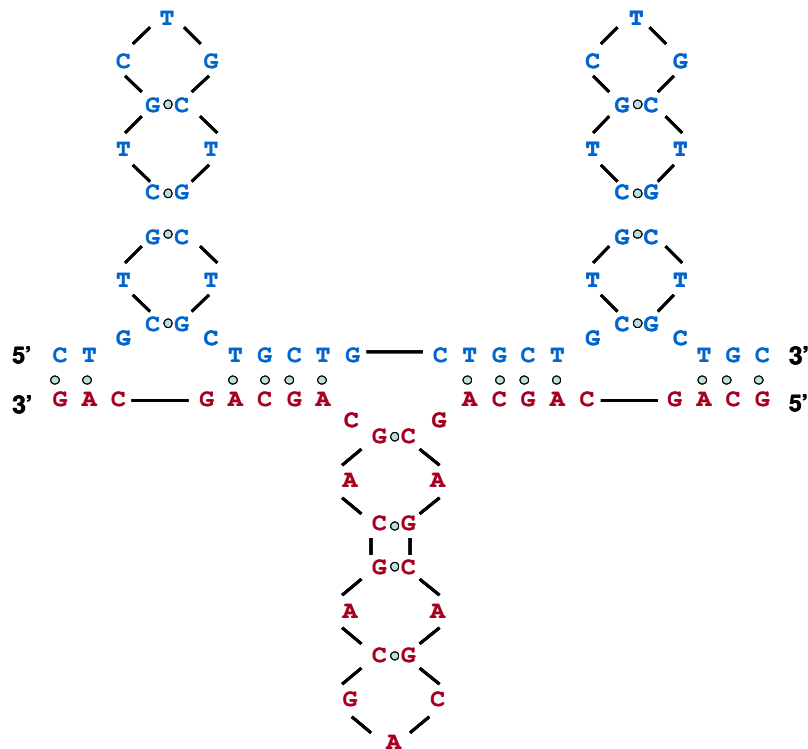


Figure 1.5 Slipped-stranded DNA structures that can be formed by expandable repeats. This diagram represents slipped-stranded structures formed by CTG • CAG tracts. In reality these TRS can consist of thousands of nucleotides that can form very large and flexible secondary structures.

Recent studies have attempted to elucidate the potential mechanisms for increases in repeat tracts. It is likely expansions occur during DNA replication as rapid accumulation of repetitive DNA requires the synthesis of large amounts of DNA and also because a portion of the lagging-strand template, known as the Okazaki initiation zone, is single-stranded and so may fold into unusual secondary structures (Mirkin, 2006; Mirkin, 2007). TRS instability commonly occurs in rapidly dividing cells or highly proliferative tissues, but are relatively stable in differentiated cells (Wohrle et al., 1995), supporting a direct role for replication in expansion. Stalling of various DNA polymerases by stable secondary DNA structures may lead to misalignment of repeats that can cause expansion or deletion events as coordination between lagging-strand and leading-strand synthesis can be disrupted (Usdin and Woodford,

1995). If synthesis continues on the leading-strand while lagging-strand synthesis resumes after skipping 1 or more Okazaki fragments due to stalling, gaps in the new lagging strand will arise that must be filled in. A contraction of repeating DNA can occur if, as DNA polymerase repairs the gap, it skips the structural element portion of the lagging-strand template, leading to loss of repeats. Alternatively, stalling in a repeat tract can initiate replication fork reversal, creating a ‘chicken-foot’ structure with single-stranded repeats in the new leading strand. This ssDNA could fold back on itself into a hairpin structure and induce repeat expansion in newly synthesised DNA (Figure 1.6) (Mirkin, 2007).

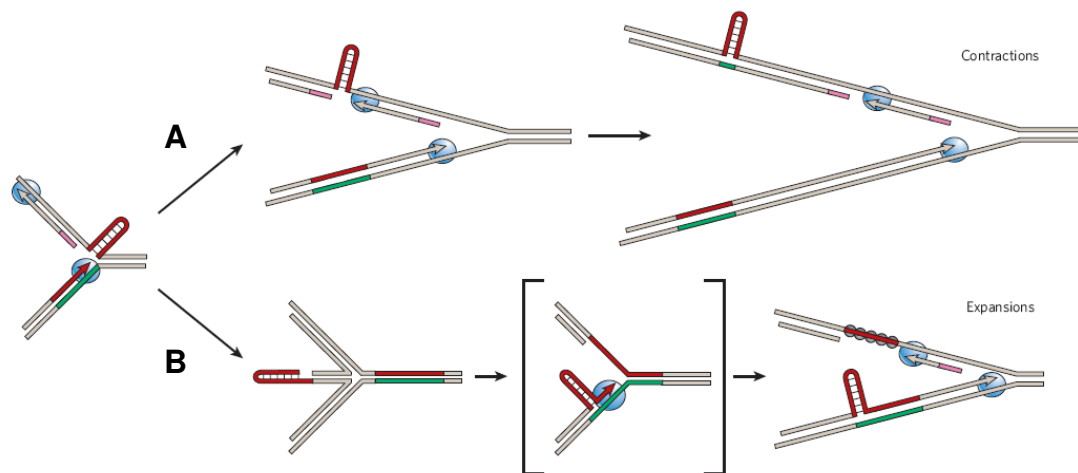


Figure 1.6 Potential mechanisms for repeat contraction and expansion. These mechanisms depend on replication fork stalling and restarting within TRS tracts. TRS (red), complementary DNA (green), Okazaki primers (pink) and DNA polymerase (blue) are shown. **A.** This pathway demonstrates how contractions may occur when lagging-strand synthesis replication machinery skips a secondary hairpin on the lagging-strand template. **B.** Expansions in TRS tracts may occur during replication fork reversal and restarting. A hairpin structure can then form on nascent leading-strand DNA as a consequence. Increases in expansions can result in larger hairpin structures. Image taken from (Mirkin, 2007).

Although the replication slippage model is considered to be a major mechanism of repeat instability, sequences capable of adopting non-B structures are also able to induce instability in non-proliferative cells, such as CAG repeat expansions in the brain (Hashida et al., 1997). This indicates a replication-independent mechanism, such as DNA repair can also induce DNA structure-related instability (Kovtun and McMurray, 2008).

Unusual secondary structures alter overall DNA conformation and, as a result, directly affect its biological behaviour by influencing processes involving interactions with DNA. Therefore, spontaneous formation of unusual secondary structures often significantly alter cellular properties and compromise cellular integrity. As discussed above, genomic stability

is not only crucial for survival of a cell, but it can also be crucial for survival of the entire organism. Global repair mechanisms may recognise certain structures as damaged DNA and, if the damage is deemed too extensive, then mechanisms will be triggered which will lead to apoptosis of the cell in multi-cellular organisms. If too many cells of a particular cell type are destroyed, the function of a particular organ essential for organism survival may be compromised.

1.3.4 DNA Repair Pathways.

As described, the various DNA-damage inducing agents and errors that occur during cellular metabolism can create dangerous base modifications, abasic sites, helix-distorting adducts, breaks in DNA and unusual structural conformations. These lesions can introduce further detrimental features such as transversions, base transitions, frameshift mutations or gross chromosomal aberrations (Kulkarni and Wilson, 2008). To protect against DNA-damage, cells contain proteins that specifically recognise and remove damaged DNA to prevent cellular dysfunction and help restore genomic stability.

Single bases can be chemically altered by modifications such as deamination and alkylation. The resulting lesions can be very mutagenic as they can cause mis-incorporation of incorrect nucleotides during DNA replication and transcription. To counteract this, cells have evolved a pathway termed base excision repair (BER) for removal of damaged bases. BER consists of 2 sub-pathways: short-patch BER, where a single nucleotide is replaced, and long-patch BER, where 2-13 nucleotides are replaced (Hakem, 2008). For both pathways, the initial step is removal of the damaged base by a specific DNA glycosylase, which occurs following DNA backbone compression and repositioning of the damaged base so that it becomes accessible for the enzyme (Berti and McCann, 2006). In mammalian cells, strand incision by APE1 endonuclease is followed by gap filling by DNA polymerase- β (short-patch) or DNA polymerase- β , ϵ or δ (long-patch) and, finally, ligation by DNA ligases I or III. Long-patch BER requires the participation of other proteins during removal and replacement of several nucleotides. Proteins involved in bacterial BER are similar to those in eukaryotic BER.

Nucleotide excision repair (NER) is an important mechanism for removing damaged nucleotides, particularly those damaged by UV light, through the activity of proteins that recognise and remove bulky distortions in the DNA double helix. NER is comprised of several subpathways of which 2 have been best characterised: global genomic-NER (GG-NER) and transcription-coupled NER (TC-NER). The only difference between these 2 pathways regards DNA lesion recognition as GG-NER repairs genome-wide damage and TC-

NER specifically repairs damage on the coding strand that inhibits RNA polymerase (Shuck et al., 2008). In humans, XPC (during GG-NER) or RNA polymerase II (during TC-NER) recognises bulky distortions in duplex DNA. This recognition stimulates DNA unwinding by transcription factor II H (TFIIH) and other cofactors. XPA, XPG and RPA stabilise the resulting complex prior to incisions 3' (by XPG) and 5' (by XPF) of the damage, and the resulting gap is filled by DNA polymerase and DNA ligase activity (Fleck and Nielsen, 2004; Hoeijmakers, 2001). Bacteria also possess both GG-NER and TC-NER, with a complex consisting of UvrA and UvrB responsible for the detection and verification in both pathways. The UvrB-DNA complex is bound by UvrC which causes incisions either side of the damaged nucleotide allowing a separate UvrD protein to subsequently remove the damaged lesion and DNA polymerase I and a DNA ligase complete the repair process. During TC-NER, the multifunctional Mfd transcription-repair coupling factor overcomes the inhibitory effect of the stalled RNA polymerase to enhance the rate of lesion repair on the template DNA strand (Savery, 2007).

Mismatch repair (MMR) is central to the repair of base insertions, deletions and misincorporation of bases that arise during DNA replication and recombination. Defects in MMR often result in alterations in DNA sequences that induce changes in amino acid sequences of peptides and affect protein function which can significantly affect cell viability (Iyer et al., 2006). Although not directly part of the MMR process, the 3'-5' exonuclease proofreading activity of DNA polymerase corrects most mismatches, but errors that are not fixed can be repaired by MMR. The importance of MMR is evident as mutation of certain MMR genes in humans can cause severe disorders, including hereditary non-polyposis colorectal cancer (Hakem, 2008). MMR is bidirectional, allowing degradation of the mismatch in either a 5' or 3' direction. In bacteria, the reaction begins by recognition of a mismatch by a MutS homodimer, followed by ATP binding to the MutS dimer and recruitment of MutL to the complex. ATP hydrolysis then induces the MutS-MutL complex to translocate along the DNA and leads to subsequent activation of MutH, which causes an incision in the non-methylated daughter strand DNA. MutS-MutL can then activate DNA helicase II to initiate unwinding of the DNA double helix, producing a single-strand of DNA containing the mismatch. Exonucleases then remove this single-stranded region and DNA polymerase III and DNA ligase complete the repair process (Jun et al., 2006). In humans, mismatches or modified bases are recognised by a MutS α heterodimer, which is comprised of MSH2 and MSH6, or a MutS β heterodimer, which is a complex of MSH2 and MSH3. Both complexes have different substrate specificities and can induce varied downstream responses. Recognition of DNA lesions primarily occurs via MutS α (Umar et al., 1994) according to a 'molecular switch model' (Christmann et al., 2003) and, once bound to a lesion, MutS α -ADP

stimulates ADP to ATP transition and intrinsic ATPase activity. This in turn induces a conformational change that permits passive sliding and association with the heterodimer MutL α complex. The DNA strand that contains the mismatch is then excised by exonuclease I (Genschel et al., 2002) followed by DNA synthesis by DNA polymerase δ (Longley et al., 1997).

1.3.5 Double-strand Break Repair.

A single inappropriate DSB poses a significant threat to genomic stability and so the efficient recognition and removal of such aberrations is of the highest priority. Cells have evolved different mechanisms to repair DSBs depending on the cell cycle phase and type of break. When a homologous DNA template is available, homologous recombination (HR) can be performed, resulting in accurate repair of DSBs. Other processes, such as non-homologous end joining (NHEJ) and microhomology-mediated end joining, are error-prone repair mechanisms that can lead to alterations in the genetic material following repair. Choice of DSB repair mechanism is also species-dependent with over 90% of DSBs in mammalian cells repaired by NHEJ, but in both yeast and bacteria, HR is the preferred repair mechanism (Hakem, 2008). The type of DSB sustained also dictates which repair pathway will be selected. An endonuclease induced DSB occurring in G1 phase of the cell cycle would be repaired through the NHEJ (Zierhut and Diffley, 2008), however ionising radiation induced damage would be preferentially repaired by HR (Barlow et al., 2008), with repair delayed if no homologous template is available.

In humans, DSB recognition is performed by ataxia telangiectasia mutated (ATM), ataxia telangiectasia related (ATR) and the DNA protein kinase catalytic subunit (DNA-PKcs). ATM and ATR attach to damaged DNA ends, which leads to kinase domain activation and subsequent activation of proteins involved in cell cycle regulation and DNA repair that share a Ser-Thr-Gln-Glu consensus sequence (Kim et al., 1999). Generally, ATM/ATR-mediated phosphorylation of target proteins such as Chk1, Chk2 and p53 is crucial in controlling G1/S and G2/M cell cycle arrest, initiation of DNA repair and apoptosis (Christmann et al., 2003).

HR is a highly accurate process that uses a homologous DNA template to repair breaks during the S or G2 phases of the eukaryotic cell cycle. In this repair pathway, the first proteins recruited to a DSB is the MRX/MRN complex, consisting of Mre11-Rad50-Xrs2 (MRX) in yeast and MRE11-Rad50-NBS1 (MRN) in humans (Bernstein and Rothstein, 2009). These complexes perform the initial processing of damaged DNA ends by producing 3' ssDNA at each end of the break. RAD51 is then recruited, which binds to the 3' single-stranded ends

and invades a homologous dsDNA, usually from a sister chromosome, forming a Holliday junction following polymerase activity. Separation of repaired DNA is achieved by resolvase enzymes and the repair is completed by ligation of the broken strands (Fleck and Nielsen, 2004).

In contrast to the accurate repair of DSBs by HR, NHEJ is error-prone and can lead to dramatic genetic alterations (Pitcher et al., 2005). During processing and ligation of the damaged DNA termini, nucleotides can be lost and the integrity of the genome may be significantly reduced. Initially it was believed NHEJ was restricted to eukaryotes, but evolutionarily NHEJ related mechanisms in prokaryotes have also been identified (Doherty et al., 2001). Unlike HR, which requires a homologous template, NHEJ can progress independent of homology between DNA duplexes, however, microhomology alignment of DNA ends is thought to facilitate the process. The NHEJ reaction requires tight coordination between several key proteins and many of these interactions are yet to be fully understood (Weterings and Chen, 2008). In eukaryotes, NHEJ is initiated by a Ku70/Ku80 dimer that recognises and binds in a ring-like structure to a variety of DSBs. In mammalian cells, once bound, Ku70/Ku80 is believed to act as a scaffold that recruits and activates DNA-PKcs and the DNA ligase IV/XRCC4 complex. DNA-PKcs has many roles, including regulation of the synaptic complex, which brings both DNA ends together and leads to activation of other proteins integral to NHEJ. For the production of ligatable ends, the DNA may need to be processed by factors such as polymerases and nucleases. It is during this processing that changes to nucleotide sequence can occur that can threaten genomic integrity (Weterings and Chen, 2008). These processing proteins are usually recruited via direct interactions with Ku or DNA-PKcs (Hoeijmakers, 2001; Pitcher et al., 2007). Once ends have been processed, ligase IV/XRCC4 joins the 2 ends together, facilitated by other proteins including XLF/Cernunnos.

The MRX complex is involved in both HR and NHEJ and its activity may determine whether repair proceeds via HR or NHEJ, at least in yeast (Daley et al., 2005). It is still unknown how HR and NHEJ are regulated in bacteria (Pitcher et al., 2007), however cell cycle status or regulation of the HR pathway by NHEJ proteins are likely to be important.

1.4 Protein Components of Non-homologous End Joining.

1.4.1 Activities of Ku Proteins.

Ku proteins were originally discovered as targets of autoantibodies in patients with polymyositis-scleroderma overlap (Mimori et al., 1981). They are now recognised as essential proteins associated with various cellular processes including NHEJ, retrotransposition and telomere maintenance (Friedl, 2002). Ku is a highly abundant nuclear protein present in all eukaryotes that is likely to be one of the first proteins to interact with DNA ends following a DSB. As mentioned, Ku are integral in detecting and interacting with DSBs and they have also been implicated in alignment and synapse formation of DNA termini. The eukaryotic Ku70/80 heterodimer exhibits high affinity binding for dsDNA ends and also to hairpin structures (Falzon et al., 1993). This affinity for DNA ends does not seem to be dependent on sequence or overhangs, allowing Ku to bind to a variety of DSB that result from DNA damage (Doherty et al., 2001). Such flexibility in substrate binding is thought to occur as Ku makes no contacts with DNA bases and only a few contacts with the DNA sugar-phosphate backbone (Riha et al., 2006). Ku binding is dependent on the DNA substrate length and it has been suggested that it can passively translocate along DNA, allowing multiple Ku proteins to bind to the same DNA molecule like beads on a string (Dynam and Yoo, 1998). This ability of Ku to translocate inwards away from the DNA end may be crucial in allowing the efficient processing and subsequent ligation of DNA termini during NHEJ. Binding affinity of Ku proteins to ssDNA is believed to be low, but there is evidence that Ku may have some binding affinity for RNA and, in some cases, binding affinity for small RNA molecules equals that of dsDNA ends (Yoo and Dynam, 1998).

Following DSB formation, DNA termini are at risk of excessive nucleolytic degradation, which could lead to the loss of important genetic information. Ku has been implicated in protection of the resulting DNA ends from this degradation as studies investigating Chinese hamster ovary cell lines lacking Ku identified deletions of hundreds of nucleotides around DSB sites (Liang and Jasin, 1996). Alongside DSB protection, Ku facilitates ligation of breaks through juxtaposition of both DNA strands and recruitment of DNA ligase IV/XRCC4 by direct protein-protein interactions (Kysela et al., 2003).

X-ray crystallographic structures demonstrated that although Ku70 and Ku80 do not share high sequence homology, the overall folding of each protein during DNA binding is similar (Walker et al., 2001). Both Ku70 and Ku80 share a similar three-domain topology - an α/β -domain at the amino-terminus, a central β -barrel domain and an α -helical arm at the carboxyl-terminus (Figure 1.7). The amino-terminal α/β -domains and helical carboxyl-terminus are

located on the Ku complex periphery and are thought not to significantly contribute to heterodimer formation or DNA binding. Instead, these regions are essential for interactions with other DNA repair proteins (Cary et al., 1997; Weller et al., 2002). The eukaryotic heterodimer forms a unique structure containing a hollow preformed channel in the centre of the ring-like structure. This region is termed the bridge and pillar region and is where DNA can thread through (Lehman et al., 2008). DNA binding is achieved using direct interactions between positively charged and hydrophobic amino acids on both Ku70 and Ku80 and the DNA sugar-phosphate backbone. Recent studies have suggested conformational changes occur to Ku once bound to DNA, which were previously unknown (Lehman et al., 2008). Other structural features of Ku include vonWillebrand factor A (vWA) domains on both Ku70 and Ku80 that become positioned near the DNA break, and 6 α -helices on the Ku80 carboxyl-terminus to which DNA-PKcs is thought to bind (Zhang et al., 2004). The vWA is involved in protein-protein interactions and is thought to be crucial in recruiting other proteins to DSBs. Other structural features of eukaryotic Ku include a SAF-A/B Acinus and PIAS (SAP) domain (carboxyl-terminus of Ku70), a domain implicated in protein-DNA interactions. SAP motifs contain a distribution of hydrophobic, polar and bulky amino acids and several positions rich in positively charged amino acids that contact the DNA backbone (Aravind and Koonin, 2000). The SAP domain is also a distant relative of the Myb DNA binding domain, which is a common feature of proteins implicated in telomere maintenance (Riha et al., 2006).

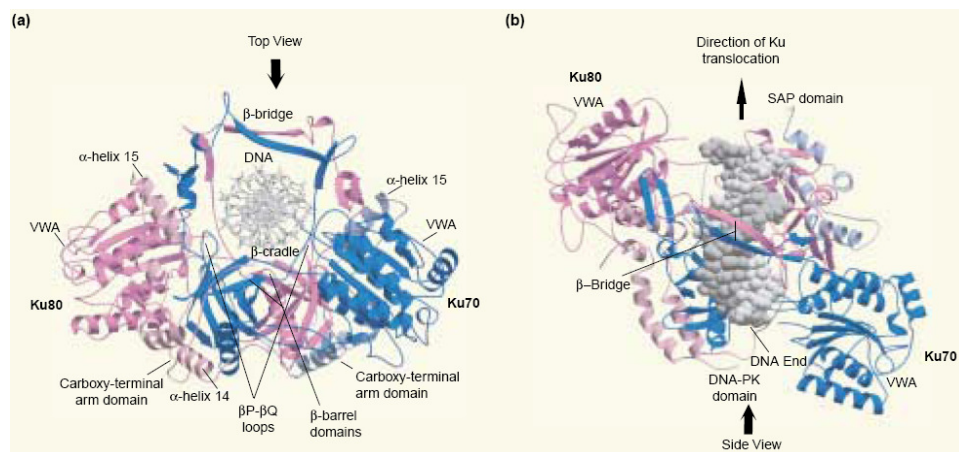


Figure 1.6 Structure of human Ku70/Ku80 bound to DNA. Ku70 is blue, Ku80 is pink and DNA is grey. **A.** Side view of Ku-DNA complex viewed down the DNA helical axis. **B.** Ku-DNA complex viewed from the top. A narrow bridge formed by 2 strands that extends across the DNA can be seen. The top of the DNA is exposed which may allow access for other proteins such as DNA ligase IV to bind directly. Ku70 SAP and Ku80 DNA-PKcs binding domain relative positions are indicated. Image taken from (Doherty and Jackson, 2001).

As well as its essential function during NHEJ, Ku plays a pivotal role in the maintenance of telomeres. As mentioned in Section 1.2, chromosome ends of linear genomes must be

excluded from DSB repair to prevent catastrophic chromosomal rearrangements. Telomere-specialised architecture, consisting of extended G-rich repeat arrays bound by specific telomeric DNA binding proteins, form higher-order non-nucleosomal structures that act as protective caps that allow telomeres to be distinguished from damage-induced DSBs (Riha et al., 2006). In most eukaryotes, the G-rich repeat forms a 3' single-stranded overhang that mediates formation of a complex t-loop structure. It is known that Ku70/Ku80 contributes to telomere end protection but the precise mode of interaction has yet to be fully elucidated. Interactions may be achieved via protein-protein interactions or by direct Ku-DNA binding. Ku deficiency can lead to a range of telomeric phenotypes in different organisms, which suggests there is no uniform model for Ku activity at telomeres. Ku has been found to protect telomeres from interchromosomal recombination and degradation, regulate telomerase activity and also aid telomere tethering to the nuclear periphery (Wang et al., 2009b).

Even though Ku appears to have important conserved roles across many species, it is only in humans where Ku is found to be an essential protein (Li et al., 2002). It was recently identified that this essential function is because of the role of Ku in telomere formation, rather than its activity during NHEJ (Wang et al., 2009b). Currently there is still no proof of direct Ku-telomere DNA binding, however, this activity would provide mechanistic explanations for the diverse functions Ku is believed to perform at telomeres.

1.4.2 Current Understanding of Prokaryotic NHEJ and Ku.

Evidence for NHEJ in prokaryotes was first obtained from the discoveries of genes in bacteria that were homologous to those involved in eukaryotic NHEJ. In contrast to eukaryotic NHEJ involving several proteins with specific roles, the best characterised bacterial NHEJ is from mycobacteria and consists of a two-component system that contains the break recognition, end-processing and ligation activity needed for NHEJ (Della et al., 2004; Weller et al., 2002). This system consists of Ku proteins that have DNA end recognition and binding ability and an associated ligase (LigD) that contains domains associated with end-processing and ligation activities. As with eukaryotic NHEJ, the bacterial NHEJ complex has the ability to synapse 2 independent DNA ends, process the termini and ligate the two strands together. There is evidence that suggests microhomologies between DNA molecules can facilitate bacterial NHEJ (Della et al., 2004) as identified with eukaryotic NHEJ, although such similarities between DNAs are not essential. A generalised schematic of NHEJ is shown in Figure 1.8.

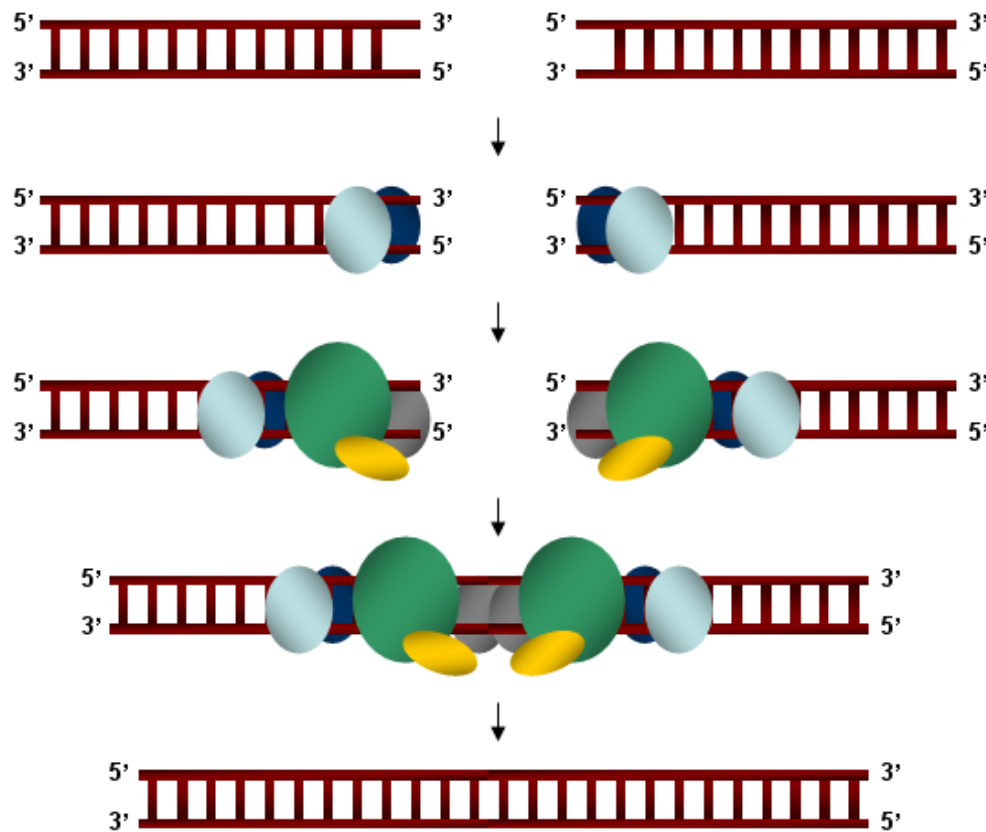


Figure 1.8 A generalised view of DSB repair by the NHEJ complex. For initiation of DSB repair, Ku dimers (blue) recognise and bind to DSB, where they may function as end-bridging and DNA alignment factors. Ku may then translocate along the DNA and allow recruitment of additional processing enzymes to the break site. Often DNA ends are not complementary and DNA end-processing, including nucleolytic activities (yellow) and gap-filling (grey), may be required to generate termini capable of ligation by DNA ligase (green). Once DNA ends have been joined, the complexes dissociate, however the mode of Ku dissociation remains unclear.

The bacterial NHEJ mechanism is yet to be fully understood. It is also unknown how much variation exists in NHEJ between different prokaryotes. Many prokaryotes that demonstrate NHEJ machinery are known to exist predominantly in a non-dividing phase of their life cycles when DNA replication is reduced or absent (Wilson et al., 2003). For example, in *Bacillus subtilis*, expression of *ykoU* (LigD) and *ykoV* (Ku) is controlled by the forespore specific sigma factor for RNA polymerase σ^G , so that expression is only induced during spore development (Wang et al., 2006). During this time, cells may be exposed to various DNA damaging agents and other factors such as desiccation that could induce DSBs. In bacteria with only 1 chromosome, the absence of a homologous DNA template means the evolution of NHEJ pathways may have been important for many prokaryotic species (Bowater and Doherty, 2006). The error-prone nature of NHEJ may also have facilitated with increased genetic diversity, which can ultimately provide selective advantages under stressful conditions by generating mutations during stationary-phase (Heidenreich et al., 2003). The

observation that several bacteria that chronically infect eukaryotic hosts contain *ku* and *ligD* homologues has lead to suggestions that NHEJ may be important during infection for repair of DSBs induced by the hosts genotoxic defence or by increasing the mutation frequency by error-prone repair (Kobayashi et al., 2008).

In eukaryotic NHEJ, it is thought formation of a synapse between the 2 DNA strands is mainly regulated by DNA-PKcs with cooperation by Ku70/80 (Riha et al., 2006). However, a degree of uncertainty remains over which protein component is integral to DNA alignment. Initially, studies indicated that Ku alone could synapse 2 independent DNA molecules (Cary et al., 1997; Ramsden and Gellert, 1998) but later investigations established that DNA-PKcs was integral to this process (DeFazio et al., 2002; Lees-Miller and Meek, 2003). Although DNA synapsis during NHEJ in eukaryotes may remain cryptic, it is clear that synapsis in prokaryotes must be DNA-PKcs independent as no homologous protein has been detected in any prokaryote. Feasibly, prokaryotic Ku may perform this activity alone as previous studies on eukaryotic Ku have alluded to, or alternatively, other NHEJ components may contribute to alignment. Studies of *Mycobacterium tuberculosis* recombinant NHEJ proteins have shown that the polymerase domain of the LigD protein (which comprises the two-component system alongside Ku) contains a functionally analogous structural element found in other NHEJ polymerases. This element is thought to facilitate connectivity between 2 discontinuous DNA ends (Juarez et al., 2006; Pitcher et al., 2007) and indicates this protein may participate in formation of a synaptic complex. It has been proposed that Ku may facilitate synaptic complex formation alongside regions of microhomology between DNAs.

Unlike the heterodimeric eukaryotic Ku that is produced from 2 separate genes, prokaryotic Ku orthologues have been characterised that are encoded by single gene and subsequent biochemical studies revealed this protein could bind as a homodimer to DSBs (Doherty et al., 2001). In addition, a number of bacteria also encode Ku proteins that have potential to form heterodimers (Pitcher et al., 2005), such as *Mesorhizobium loti*, *Streptomyces coelicolor* and *Sinorhizobium meliloti* (Kobayashi et al., 2008), as they contain multiple genes with homology to *ku*, indicating that NHEJ may show variation between different prokaryotic organisms. Interestingly, bacterial *ku* genes are often genetically linked to operons containing an ATP-dependent DNA ligase gene, which together may form the two-component system. These Ku-associated ligases frequently possess domains that demonstrate sequence homology with eukaryotic nucleases and polymerases and could explain how bacterial NHEJ compensates for absent end processing factors that are found in eukaryotic NHEJ (Weller and Doherty, 2001). Although *ku* homologues have been characterised in bacteria, certain strains such as *Escherichia coli* K12 (enterobacteria) show no evidence of *ku* related genes. This lack of any

obvious phylogenetic pattern between bacteria with or without NHEJ apparatus indicates these genes are likely to have been passed between organisms through horizontal gene-transfer (Bowater and Doherty, 2006).

The bacterial Ku homologues are significantly smaller (30-40 kDa) than their eukaryotic counterparts and generally only contain domains similar to the central core DNA binding and dimerisation region of Ku70/Ku80 (Figure 1.9). They also usually lack other domains associated with eukaryotic Ku proteins, including the SAP (located on the carboxyl-terminal domain (CTD) of Ku70) and the vWA domain (located on the amino-terminal domain (NTD) extension of Ku70/Ku80) (Pitcher et al., 2007), which are involved in protein-DNA and protein-protein interactions respectively. Structural information regarding bacterial Ku binding has not been fully identified, but sequence analysis predicts these proteins form dimeric complexes that are structurally related to eukaryotic Ku, with a central ring-like structure. These similarities suggest prokaryotic and eukaryotic Ku derived from the same common ancestor such as the phage protein Gam (Figure 1.10).

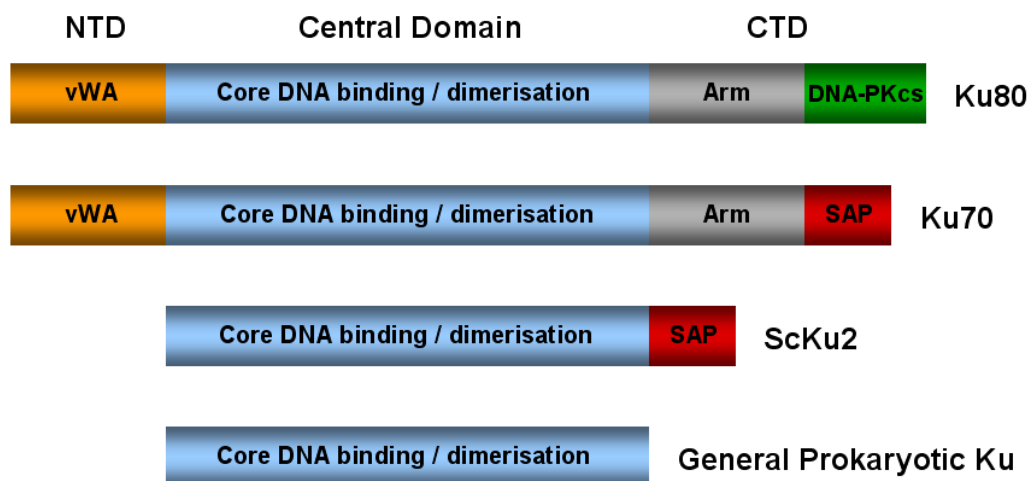


Figure 1.9 Domain organisation of eukaryotic and prokaryotic Ku proteins. Eukaryotic Ku proteins are generally larger and contain additional specialised domains in comparison to their prokaryotic counterparts. Ku80 is the largest protein, with a vWA protein-protein interaction domain on the NTD, a central DNA binding / dimerisation domain and a CTD containing a helical arm region and also a domain implicated in interactions with DNA-PKcs. Ku70 has similar NTD and core regions but the CTD differs as a SAP DNA-binding domain is present instead of the DNA-PKcs interacting domain. Prokaryotic Ku are considerably smaller and most only consist of the core DNA binding / dimerisation domain except for ScKu2 which contains a domain homologous to the SAP domain of Ku70 on the CTD.

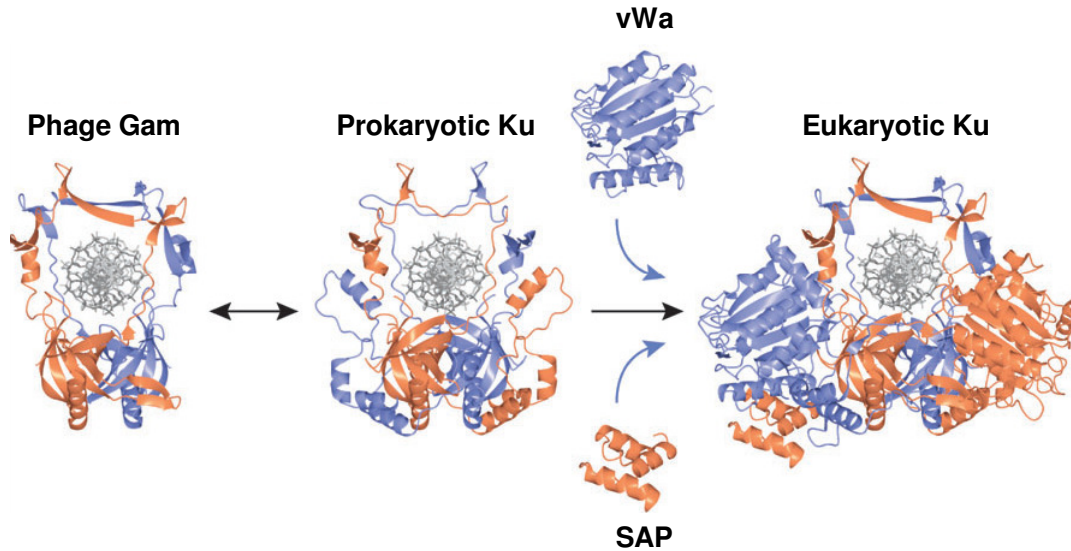


Figure 1.10 Evolution of Ku proteins. As Ku proteins can be found in all kingdoms it is believed their existence predates branching of prokaryotic and eukaryotic domains. Primordial Ku may have been a homodimeric protein that demonstrates significant similarities to phage Ku-like proteins (Gam) that act as linear DNA end protection proteins. Interactions with DNA ligases led to the two-component NHEJ seen in some prokaryotes and gene duplication may have produced heterodimeric prokaryotic Ku. Additional vWA and SAP domains incorporation resulted in the modular architecture and functional characteristics seen in the Ku70/Ku80 eukaryotic Ku. Taken from (Pitcher et al., 2007).

Aside from NHEJ, alternative roles of prokaryotic Ku remain unknown. Investigations of Ku from *M. tuberculosis* imply prokaryotic Ku contains similar DNA-binding characteristics as eukaryotic Ku, with a preference for dsDNA ends and the ability to load onto DNA dependent on DNA length, suggesting the ability to translocate along DNA (Weller et al., 2002). Eukaryotic Ku is heavily involved in telomere regulation, however no evidence exists yet for DNA end maintenance by prokaryotic Ku. Although the current activity of Ku proteins in prokaryotes may now be confined to NHEJ, in the past Ku may have been instrumental in prokaryotic evolution. It has been determined that bacteria expressing Ku homologues are more adept at acquiring genetic material from their surrounding environment and incorporating this into their own genomes (d'Adda di Fagagna et al., 2003), a feature that would have greatly assisted the genetic diversity now evident in nature.

1.4.3 Ku proteins of *Streptomyces coelicolor*.

Streptomyces coelicolor are Gram-positive bacteria that derive from the largest genus of Actinobacteria. *S. coelicolor* are soil dwelling and have complex lifecycles, involving multicellular entities with biochemically and morphologically differentiated cell types, with each individual cell type comprising a specific function (Figure 1.11) (Miguel et al., 2000).

S. coelicolor are perhaps best known for their complex secondary metabolism and their involvement in production of several antibiotics used today (Paradkar et al., 2003).

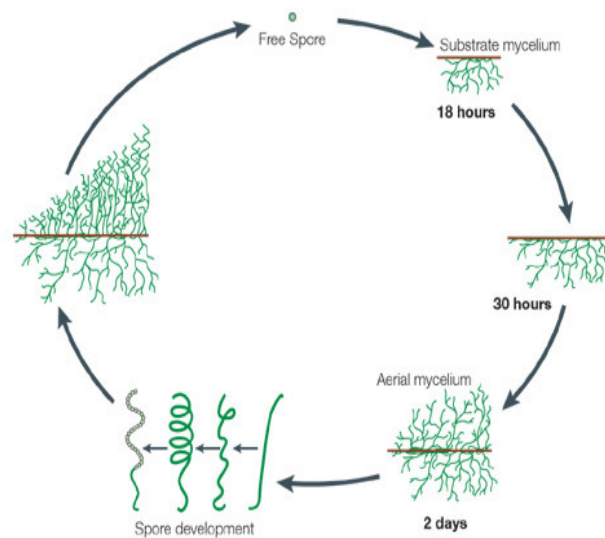


Figure 1.11 The complex lifecycle of *S. coelicolor*. Entities at each individual stage are biochemically and morphologically different and have functions specific for each particular stage. Taken from (Angert, 2005).

Colonies of *S. coelicolor* grow vegetatively by the formation of multigenomic hyphae that form branching substrate mycelium (Gehring et al., 2004). Under stressful conditions such as nutrient deprivation, aerial mycelium filaments are produced that are capable of penetrating the surface tension at the soil interface. The aerial filaments can differentiate into spore-bearing structures capable of dispersal (Tyers and Buttner, 2004). Prior to sporulation, regularly spaced sporulation septa divide hyphae into unigenomic prespore units that mature during sporulation and are released as free spores that are hydrophobic dormant bodies (Hirsch and Ensign, 1976). For germination of the spores, exogenous nutrients including water are required (Ensign, 1978).

Recently, 3 Ku homologues have been identified in *S. coelicolor*, known as SCO5309 (1095 nucleotides), SCO0601 (1041 nucleotides), which are both chromosomal, and SCP1.285c (897 nucleotides), which is located on plasmid SCP1. SCO5309 and SCO0601 are predicted to have molecular weights of approximately 39 kDa and SCP1.285c is approximately 33 kDa (Sayer, 2008). All 3 Ku homologues demonstrate sequence similarities with the core region of eukaryotic Ku70/Ku80. SCO0601 also contains a helix-extended region-helix (HEH) domain on the carboxyl-terminal extension, which is very similar to the SAP domain found on Ku70. As described in Section 1.4.1, SAP domains contain positively charged amino

acids that interact with negatively charged helical surfaces. They consist of a short mobile helix adjacent to a second helix, which is separate from a third helix that is parallel to the second helix (Aravind and Koonin, 2001). It is likely that SAP domains evolved from HEH domains and both are thought to have similar properties. Apart from sequence homology with Ku70/Ku80, which suggests a NHEJ mechanism exists in *S. coelicolor*, little else is known regarding the specific activity or other specialised functions for these proteins. Interestingly, SCO5309 is located adjacent to a gene that shows high sequence homology to the primase region of the *mtligD* ligase implicated in NHEJ in *M. tuberculosis* (Weller and Doherty, 2001).

S. coelicolor is unlike many other prokaryotes in that it contains a linear genome and can also contain linear plasmids that demonstrate the ability to rapidly circularise. These properties have made *S. coelicolor* an interesting experimental system for the investigation of telomere characteristics. *Streptomyces* replicons are dsDNA molecules with their 5' ends covalently attached to terminal proteins (Bao and Cohen, 2001; Hirochika and Sakaguchi, 1982; Kinashi and Shimaji, 1987; Yang et al., 2002). Replication of *Streptomyces* linear plasmids proceeds divergently from a central site in the DNA and generates 3'-leading strand overhangs at the telomeres (Bao and Cohen, 2003). Like eukaryotic telomeric regions, *S. coelicolor* chromosomal ends contain repeating sequences, however, unlike eukaryotic tandem direct repeats, *S. coelicolor* repeats contain inverted sequences. The chromosomal arms are highly unstable and large chromosomal rearrangements can occur within these regions. Consequently, no essential genes are located within 1.2 Mb of either end of the *S. coelicolor* chromosome with the exception of *argG* (Redenbach et al., 1996). Telomeres of *S. coelicolor* range in size, show high instability and telomeric DNA is often lost when replicons circularise (Qin and Cohen, 2002). Deletions also commonly occur in *S. coelicolor* telomeric regions either by replicative transposition of mobile DNA, replication fork collapse at ssDNA breaks or exonuclease degradation when terminal proteins are absent (Volff and Altenbuchner, 1998). Deletion events are also regularly associated with DNA amplifications that can both dramatically lower genomic stability.

Unlike eukaryotic Ku70/80, there is no evidence that bacterial Ku are involved in telomere maintenance (Pitcher et al., 2007). However, as *S. coelicolor* contains linear replicons with highly unstable chromosomal ends, it would be of interest to determine whether any of the Ku homologues demonstrate telomere regulation activities that are associated with their eukaryotic counterparts.

1.4.4 Non-homologous End Joining DNA Ligase Proteins.

DNA ligases catalyse the formation of covalent bonds between adjacent 3'-hydroxyl and 5'-phosphate groups on substrate DNA. The reaction mechanism is dependent on a divalent cation (usually Mg^{2+}) and an ATP or NAD^+ cofactor that donates an adenylate group onto the DNA ligase to form a phosphoamide intermediate. The adenylate group is subsequently transferred to the 5'-phosphate, which is then subject to nucleophilic attack from the 3'-hydroxyl group, causing release of AMP and formation of a phosphodiester bond (Wilkinson et al., 2001). Most organisms contain multiple specialised DNA ligases that function only during particular cellular processes including replication, recombination and repair of DNA.

In eukaryotes, successful joining of DSBs during NHEJ is performed by the ATP-dependent DNA ligase IV in complex with XRCC4 (Hoeijmakers, 2001). Prokaryotic NHEJ requires the activity of ATP-dependent LigD proteins that are often found in operons alongside Ku homologues (Pitcher et al., 2007). As described in Section 1.4.2, these DNA ligases can be multifunctional and contain additional domains for processing DNA prior to ligation. The LigD from *Mycobacterium tuberculosis* (MtLigD) exhibits both polymerase and nuclease domains in addition to the central ligase domain whereas other prokaryotes, such as *Bacillus subtilis*, contain LigD proteins with only additional nuclease regions (Aravind and Koonin, 2001). Biochemical experiments have confirmed that these proteins have their predicted enzymatic activities (Pitcher et al 2007). Other prokaryotes contain *ligD* genes that only encode ligase domains, indicating additional proteins would be required for any necessary DNA processing in these species. The *ligD* gene of *S. coelicolor* (*scligD*) displays high sequence homology with the DNA ligase domain of LigD from *M. tuberculosis* but potential polymerase or nuclease domains are absent. Interestingly, 3 genes are present in *S. coelicolor* that show homology to the polymerase region of MtLigD and a separate gene is also present that displays homology to the MtLigD nuclease region. Analysis of ScLigD has revealed this protein is able to ligate varied DNA substrates, albeit at a relatively low rate (Sayer, 2008). Potentially, the inclusion of additional NHEJ proteins, such as ScKu, may enhance this rate to produce an efficient DSB repair pathway *in vivo*.

1.5 p53 Tumour Suppressor Protein.

1.5.1 p53 Structure and Function.

Initially identified as an oncogene, *p53* is now recognised as a gene coding for a major tumour suppressor protein that controls an extremely complex network within multicellular organisms. *p53* exerts its function by directly or indirectly regulating numerous target genes involved in essential cellular processes that directly influence the maintenance of genomic stability (Laptenko and Prives, 2006). Due to its significance, it has been termed the ‘Guardian of the genome’ and its importance is reinforced by the fact that this protein is mutated in around half of all known cancer cases (Shu et al., 2006). General *p53* activity in unstressed cells is very low. However, following induction of stress, various pathways are activated that alter post-translational modifications and the stability of *p53*, resulting in its accumulation in various stable forms that allow it to perform diverse roles throughout the cell. A well documented function of *p53* is to act as a transcription factor that binds to specific consensus sequences located in regions of the genome, with subsequent alteration of expression of a particular gene target. Numerous targets have been identified that play roles in cell cycle arrest, DNA repair, apoptosis and senescence (Appella and Anderson, 2001; Murray-Zmijewski et al., 2008; Oren, 2003; Vogelstein et al., 2000; Vousden, 2006). Some activated *p53* target genes are responsible for cell cycle arrest to prevent replication of potentially dangerous DNA, allowing time for decisions to be made as to whether the damage can be repaired or if the cell should undergo apoptosis. Other downstream genes are directly involved in removing dangerous DNA lesions by DNA repair processes, and other targets are instrumental in the induction of apoptosis. Transcription-independent activities of *p53* in the cell cytosol leading to apoptosis initiation have also been revealed, which have added to the complexity of understanding how this protein functions (Romer et al., 2006). Activation of *p53* is usually dependent upon genotoxic or cellular stress that causes damage to DNA, which can pose a significant risk to the stability of a cell. Although activation is crucial for many of the observed roles of *p53*, it has recently become apparent that basal inactivated *p53* can also influence cellular activity, including regulation of self-renewal of adult neural stem cells (Meletis et al., 2006) and repression of CD44 cell surface proteins (Godar et al., 2008). One model (Figure 1.12) has been proposed that suggests at low levels *p53* functions in response to constitutive stress associated with growth and development to regulate necessary DNA repair, but, following acute stress, *p53* activity is significantly increased and apoptosis inducing functions are activated (Vousden and Lane, 2007).

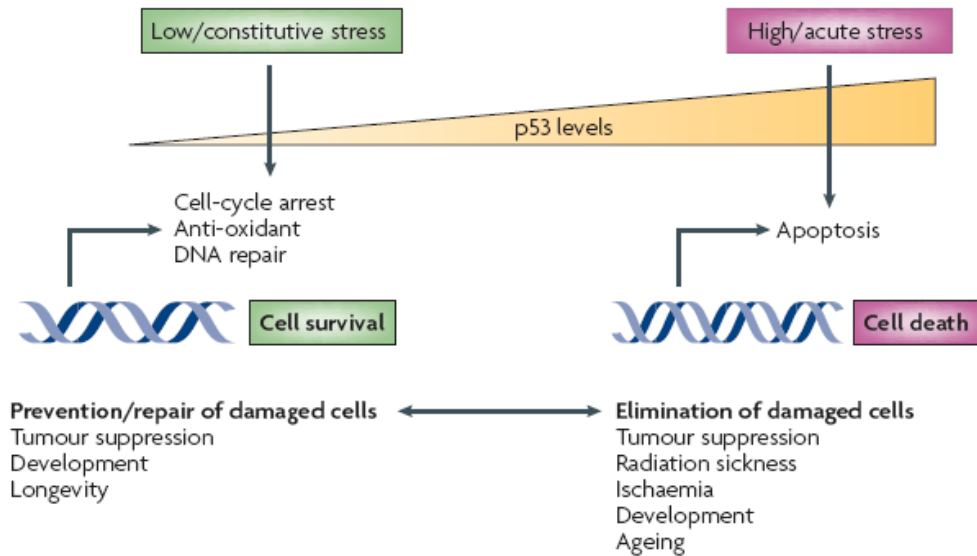


Figure 1.12 Determining the p53 response. The p53 induced response depends on the type of stress encountered. During low or constitutive stress, p53 may function to reduce oxidative damage and control repair mechanisms to restore genome stability. Such activity contributes to the general health of the cell and prevents acquisition of tumourigenic mutations. Following acute stress, p53 is fully induced and apoptotic pathways can be initiated to eliminate dangerous cells. Taken from (Vousden and Lane, 2007).

Malignant transformation of cells requires DNA damage and oncogenic signalling, 2 stresses that induce p53 activity. In non-cancerous cells, p53 will sense and react to resolve these errors, either by initiating DNA repair or cell death. If p53 function is prevented, cancerous cells will continue to divide, which can ultimately prove catastrophic for the organism. A large number of human cancers arise because of inactivation of p53 pathways, with many cancers resulting from mutations of p53 itself. The majority of these are missense mutations occurring in the core domain of the protein that is involved in binding to DNA. A vast array of different mutations have been identified in this region and only a single amino acid change can be enough to prevent p53-DNA binding ability (Brachmann, 2004). A missense mutation can lead to prevention of crucial DNA contact, loss of local structure of the core domain or even unfolding of the entire core domain. Mutant p53 can also have a dominant-negative effect by interfering with wild-type p53 proteins. Although carcinogenesis usually requires the loss of both alleles of most tumour suppressor proteins, the dominant negative effect of p53 means a single mutant allele can be enough for a cell to become cancerous (Willis et al., 2004). Mutations in germline p53 can be hereditary and this is demonstrated in individuals with Li-Fraumeni Syndrome, a rare autosomal dominant hereditary disorder characterised by mutations in the p53 gene that greatly increase susceptibility to cancer.

As a transcription factor, p53 binds to DNA as a homotetramer composed of modular domains that are crucial for its full activity. Each monomer consists of 393 amino acids organised into separate functional domains (Kitayner et al., 2006): an amino-terminal activation domain (amino acid residues 1-67), a proline-rich region (residues 67-98) a central DNA binding domain (CDBD) (residues 98-303), a nuclear localisation signal containing region (residues 303-323), a tetramerisation (TET) domain (residues 323-363) and a carboxyl terminal domain containing a basic regulatory region (363-393) (Arrowsmith, 1999; Okorokov and Orlova, 2009) (Figure 1.13).

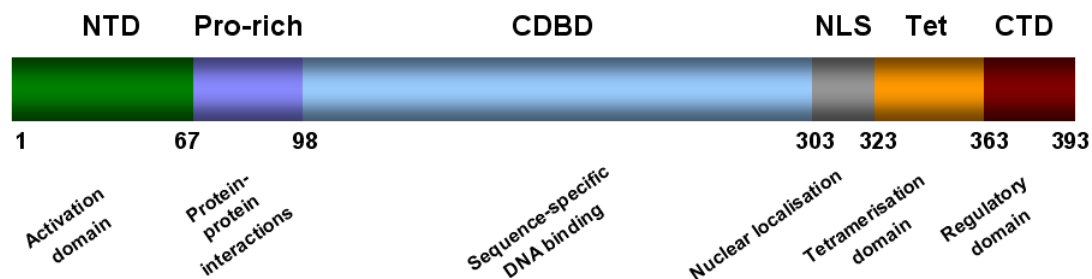


Figure 1.13 Schematic representation of human p53 domains. p53 is composed of an amino-terminal domain (NTD), proline-rich region (pro-rich), central DNA-binding domain (CDBD), nuclear localisation signal (NLS), tetramerisation domain and a regulatory carboxyl terminal domain (CTD). Each domain elicits a specific function that contributes to the diverse activity of p53.

The NTD contains a region that regulates p53 transcriptional activities and also a region rich in proline amino acids that acts as a target for proteins containing SH3 domains. This proline-rich area contains 5 copies of the amino acid sequence PXXP, which allows specific protein-protein interactions to occur in the cytosol that can instigate apoptosis (Romer et al., 2006). The NTD activation domain is an unfolded region containing islands of secondary structures that provide target binding sites for proteins that also participate in transcription and can induce expression of target genes. This domain is also subjected to negative regulation by Mouse double minute 2 (Mdm2) and adenovirus E1B proteins. Extensive post-translational modifications to this domain, such as phosphorylations, affect p53 stability and activity (Dawson et al., 2003) and complete removal of this domain has been documented as stabilising sequence-specific p53 binding (Cain et al., 2000). A recently discovered functional domain between the proline-rich and central domain of p53 has been found to contain repressing activity, but factors that interact with this domain are currently poorly understood (Curtin and Spinella, 2005).

The first of the 2 autonomous DNA binding domains are present on p53 is the CDBD. This domain spans amino acids 98-303 and is essential for binding to specific DNA sequences in

promoter regions of target genes referred to as p53 response elements (REs) (described in Section 1.5.3). More than 90% of p53 mutations that lead to carcinogenesis occur in this domain and over 1300 currently known mutations lead to partial or complete loss of binding activity (Brachmann, 2004). Some mutations eliminate direct contacts between p53 and DNA, and others destabilise the tertiary structure of p53, which prevents DNA binding (Bullock et al., 2000). This region of p53 is more stable than other regions as it has high resistance to proteolytic degradation. The CDBD contains an anti-parallel β -sheet and 2 loops that are stabilised by Zn^{2+} to form a loop-sheet-helix motif which comprises the sequence-specific DNA binding determinants of the core domain. A double salt bridge is located within the core domain and is thought to be important for cooperative binding between monomers (Dehner et al., 2005). Structural analysis of the CDBD domain bound and unbound to DNA have demonstrated the conformation remains relatively unchanged following DNA binding (Ho et al., 2006).

A tetramerisation domain is essential for controlling the quaternary structure of p53 by assisting the formation of homotetramers that are required for p53 DNA binding activity. The p53 tetramer consists of a dimer of dimers, with fully folded primary dimers but only partially formed inter-dimer interactions. The monomer contains an α -helix and β -strand linked in a sharp turn that assembles with 3 other monomers via an anti-parallel helix-helix interface held together by hydrophobic interactions (Joerger and Fersht, 2008). In solution, an equilibrium exists between p53 dimers and tetramers, but tetramerisation dramatically increases in the presence of p53 REs (Romer et al., 2006). p53's ability to perform multiple functions when in its correctly folded structure is provided by an intrinsically flexible quaternary organisation, allowing for the diverse activities associated with p53 (Hupp and Lane, 1994). A linker region containing the nuclear localisation sequence of p53 is situated between the CDBD and tetramerisation domain.

The second DNA binding domain of p53 is the carboxyl terminal regulatory domain (CTD) which is located adjacent to the tetramerisation domain. This domain binds directly to DNA through a sequence-independent manner and can regulate sequence-specific binding by the core CDBD. The CTD is unstructured and contains high levels of serine and lysine residues that allow several types of modifications to occur, such as phosphorylations and acetylations. It is believed such modifications significantly increase p53 binding affinity for its consensus sequence (Bayle et al., 1995). CTD lysine residues may also regulate p53 stability as they can be ubiquitinated by Mdm2, which can result in proteolytic degradation, but these same lysines can also act to stabilise p53 following acetylation by histone acetyl transferases.

Due to its physiological significance, it is of no surprise that p53 is heavily regulated to ensure it is only active when required. To this effect, in unstressed cells p53 is found at very low levels in its latent form. Following cellular stress signals, p53 is activated via various post-translational modifications that allow p53 levels to accumulate. There are 17 serine and threonine residues located in the NTD, CDBD and CTD that are possible phosphorylation sites for kinases such as ATM, which phosphorylates serine 15, and cyclin dependent kinases (CDKs), which phosphorylate serine 20 (Somasundaram, 2000). Phosphorylations generally decrease the turnover and increase the sequence-specific binding affinity of p53. ATR and Chk1 are able to phosphorylate multiple p53 residues following DNA damage resulting from UV light exposure (Romer et al., 2006). For a comprehensive review of p53 regulation see (Boehme and Blattner, 2009).

1.5.2 Downstream Events following p53 Activation.

Transactivation of p53 targets relies on the sequence-specific binding of p53 to consensus DNA in the promoter regions of target genes. Products encoded by these genes are involved in processes including growth arrest, DNA repair, apoptosis and senescence (Itahana et al., 2001; Romer et al., 2006; Sax and El-Deiry, 2003; Shu et al., 2006). Sequence-specific p53-DNA binding generally results in target gene up-regulation, however down-regulation of certain genes has also been documented (Li et al., 2004). It is believed p53 interactions with consensus DNA may facilitate promoter opening via recruitment of chromatin remodelling factors or histone transacetylases, thereby allowing access for general transcription factors and RNA polymerase (Horn and Vousden, 2007). Transcription can also be stimulated by p53 through direct interactions with Mediator complex components which facilitates pre-initiation complex formation and also by recruitment of basal transcription factors such as TFIIA and TFIID.

Growth arrest is the reversible prevention of cell cycle progression. Downstream p53 target p21^{WAF1/CIP1} is a CDK inhibitor that can induce cell cycle arrest in G1 phase by eventual prevention E2F activation of genes involved in DNA synthesis (Sax and El-Deiry, 2003). G1 arrest is also achieved through p53 regulated inhibition of cyclin E through increased expression of hCDC4b. p53 can induce G2 arrest by direct activation of 14-3-3 σ (Bargonetti and Manfredi, 2002) or by inhibiting cdc25c which controls cyclin B activity.

DNA repair pathways can be activated by p53 in order to restore genomic stability following DNA damage. Repair pathways known to be initiated in response to p53 activation include NER, BER and MMR (Sengupta and Harris, 2005). The role of p53 in NER is to induce

expression of genes encoding GADD45 and xeroderma pigmentosum group C and E (XPC and XPE). GADD45 can modulate chromatin and allow other repair components to access DNA, XPE localises at damaged DNA and facilitates UV photoproduct removal and XPC is involved in recognition of bulky adducts in DNA. Direct interactions between p53 and TFIIH and the helicases XPB and XPD (Harms et al., 2004) allows p53 to directly participate in both NER and BER. BER is also influenced by p53 as interactions between BER repair machinery and DNA polymerase β are stabilised by p53. Two key components of MMR that are up-regulated following p53 activation are MutS homologue 2 (hMSH2) and proliferating cell nuclear antigen (PCNA). hMSH2 is involved in mismatch recognition and binding to mismatched bases and PCNA regulates hMSH2 transfer to mismatched bases (Xu and Morris, 1999). MMR can also be regulated by p53 as expression of *p53R2*, a gene with homology to the R2 regulatory subunit of ribonucleotide reductase, is directly controlled by p53. *p53R2* increases the pool of free dNTPs to allow repair of DNA to proceed.

p53 can initiate apoptosis either transcription-dependently or independently. Several genes transcriptionally-transactivated by p53 are involved in apoptosis induction following cellular stress via both intrinsic and extrinsic apoptotic pathways (Fridman and Lowe, 2003). Genes that trigger apoptosis following transactivation by p53 include Bcl-2 family members such as Bax, Puma, Noxa and Bid, and non-Bcl-2 family protein p53-apoptosis inducing factor 1 that can all induce the intrinsic apoptotic pathway (Yu and Zhang, 2005). Transcription-independent activation of apoptosis by p53 has also been identified during the intrinsic apoptotic pathway as p53 can directly interact with Bcl-x and other Bcl-2 proteins, which leads to their inhibition and the subsequent release of cytochrome c, resulting in activation of the caspase cascade (Fridman and Lowe, 2003). Direct signalling by p53 to the mitochondria can activate Bax in the cytosol and mitochondrial p53 can induce apoptosis by activating Bak proteins. Such transcription-independent mechanisms are thought to act in conjunction with transcription-dependent mechanisms to ensure progression into programmed cell death is tightly regulated. The extrinsic apoptosis pathway is also regulated by p53 as Fas/CD95, DR5 death receptor loci and the Fas ligand genes are all up-regulated by p53 (Harms et al., 2004).

Irreversible cell-cycle arrest is a characteristic of cells that have undergone senescence. This process can be regulated by p53 to prevent cell division in cells with damaged DNA, however the mechanisms of this regulation are still unclear. It is known that Ras induced senescence is dependent on increases in cellular p53 and levels of p21^{WAF1/CIP1} are also at their highest in senescent cells (Itahana et al., 2001).

Transcriptional activation of target genes by p53 has been well established, however negative regulation of other genes has also been identified, leading to down-regulation following p53 induction. This usually occurs when a direct p53 target suppresses another gene. p53 can inhibit transcription of RNA polymerases I, II and III via different mechanisms such as repressing transcription activators, interfering with transcription machinery complexes or recruitment of chromatin modifying factors, including histone deacetylase (HDAC) (Laptenko and Prives, 2006). Transcriptional repression can also be dependent upon activation of other p53 target genes as repression of Chk1, Cyclin A2 and survivin can result from transcriptional activation of p21^{WAF1/CIP1} and subsequent cell cycle arrest.

1.5.3 p53-DNA Interactions.

Interactions between p53 and nucleic acids are highly versatile and show considerable flexibility, undoubtedly facilitating p53 to perform its diverse range of activities. The recognition and binding affinity of p53 for DNA is heavily influenced by the presence of consensus response elements in the DNA or by structural determinants conforming to non-B DNA. High-affinity non-specific DNA interactions to linear dsDNA have also been revealed that are likely to promote location of specific response elements by locating p53 to DNA (McKinney et al., 2004).

Perhaps the best understood mode of p53-DNA interactions is the binding of p53 to response elements in target gene promoter regions, leading to alterations in transcriptional activation. This type of interaction is complex and is determined by both sequence and topology of the promoter region. These DNA elements consist of 2 palindromic decamer motifs of the sequence 5'-RRRCWWGYYY-3' (where W is A or T, R is C or G and Y is C or T), usually separated by 0-13 base pairs (el-Deiry et al., 1993; Wang et al., 2009a). Recently it has been identified that genes repressed by p53 contain an alternative sequence of 5'-RRXCXXGXYY-XRXCXXGXYY-3' (Wang et al., 2009a), however this repressive activity of p53 is still to be fully elucidated. In nature, genomic p53 consensus sites rarely conform precisely to these sequences and proven sites are often significantly divergent and some sites only contain 1 half site or less (Kim and Deppert, 2006). Differences among consensus sites influence p53 binding affinity and kinetics of gene transactivation (Resnick-Silverman et al., 1998). Generally, strong p53 binding sites contain obligate C and G residues, a 0 or 1 base pair space region between half-sites and contain less than 3 non-consensus bases (Inga et al., 2002). Known p53 target sites that demonstrate low similarity to the consensus sequence are believed to contain specific architectural features that can induce recognition and subsequent interaction. Studies investigating p53 binding affinity to various consensus sequences have

demonstrated that binding affinity is highest for genes involved in cell cycle arrest and DNA repair, and lower for pro-apoptotic target genes (Weinberg et al., 2005). Therefore, it is strongly believed the different affinities of p53 to various DNA sequences modulates the activation of target genes. It is known that at low concentrations, p53 usually transactivates genes involved in cell cycle arrest or DNA repair. As p53 interacts with high affinity to promoters of genes such as *p21* or *14-3-3a*, the probability of p53 interacting with these sites is significantly greater than interactions with lower affinity promoter regions of apoptotic genes, including *Bax*. Following induction of stress, p53 becomes stabilised and levels increase, significantly increasing the chances of binding to pro-apoptotic promoter regions (Weinberg et al., 2005).

Following consensus DNA recognition, binding of p53 (Figure 1.14) causes DNA conformational changes by twisting and bending the DNA molecule by around 20° at the site of contact (Ho et al., 2006). Sites in DNA that closely conform to the consensus sequence are predicted to show higher flexibility and allow favourable protein-DNA contacts to occur that increase p53 binding affinity (Weinberg et al., 2005). DNA-binding is also dependent on the cellular environment, modifications to p53, chromatin architecture and other nuclear proteins.

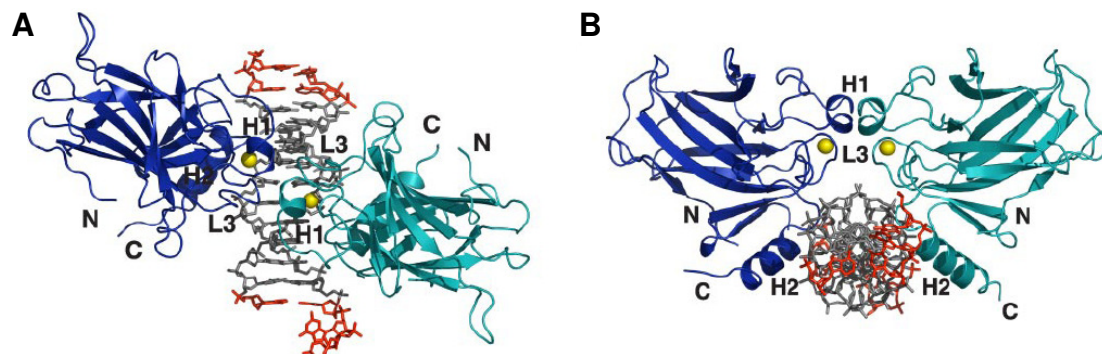


Figure 1.14 Structure of a mouse p53 DNA binding domain dimer bound to DNA. A. Top view of the overall structure of p53 DNA binding domain dimer (blue and green) bound to DNA consensus sequence (grey). Zinc ions are represented in yellow. **B.** View of the same structure but with DNA helix rotated 90°. Image taken from (Ho et al., 2006).

Post-translational modifications influence p53 binding affinity for its consensus sequence (Bayle et al., 1995). Initial studies identified that phosphorylation of serine 392 or binding of PAb421 antibody to the CTD increased DNA binding, suggesting that modifications to latent p53 CTD following cellular stress promoted p53 binding to consensus targets (Ahn and Prives, 2001). Other modifications to the CTD, such as acetylation and even complete removal of the CTD during *in vitro* studies, were also found to stimulate sequence-specific binding, which introduced speculation that the CTD negatively regulated this mode of DNA

binding and required modifications to allow p53 to bind to target sites. Models were proposed that suggested the CTD prevented sequence-specific binding either by directly interacting with the CDBD or by altering the overall tertiary structure of p53, and that post-translational modifications were required to stop this repression (Liu and Kulesz-Martin, 2001). As the CTD can also directly interact with DNA and can bind with high-affinity to non-specific DNA, this characteristic was also thought to prevent sequence-specific DNA binding as interactions with non-specific DNA would interfere with p53 binding to consensus binding sites (Anderson et al., 1997; Yakovleva et al., 2001). Generally, it was accepted the CTD required modification to switch p53 from an inactive latent form to an active sequence-specific DNA binding form (Kim and Deppert, 2006).

Recently, it was realised there is no latency regarding p53 sequence-specific binding (Kim and Deppert, 2003) and that the previous model was derived from unreliable *in vitro* investigations studying p53 binding to short-linear DNA substrates (Anderson et al., 1997; Hupp and Lane, 1994). It was also identified that sequence-specific binding was determined by DNA topology and not a conformational change in p53 itself (Gohler et al., 2002). Subsequent structural analysis identified that the CTD does not influence the CDBD conformation, casting further doubt on the latency model. Eventually it was realised that in contrast to previous models, the CTD was actually essential to p53 sequence-specific binding as deletion of the CTD *in vivo* prevented binding to consensus DNA and impaired p53-dependent transcription (McKinney et al., 2004). It was proposed that the CTD may allow p53 to bind sequence-independently to DNA and then slide along the DNA searching for p53 consensus regions that p53 binds to with higher affinity.

1.5.4 p53 and Non-B DNA.

The CTD has not only been shown to interact with non-specific DNA, but, with the cooperation of the CDBD and NTD, to also promote binding to certain types of aberrant DNA structures in a sequence-independent manner (Lee et al., 1995; Reed et al., 1995; Stansel et al., 2002). Direct recognition and interactions with damaged DNA indicates p53 may participate in DNA repair transcription-independently by acting as a platform to recruit repair factors. Tight interactions with damaged DNA may prevent p53 degradation by alterations in p53 conformation, modification or association with other proteins.

It has also been identified that p53 does not only bind to structures resulting from DNA damage, but also to unusual non-B DNA conformations including mismatches, bulges (Degtyareva et al., 2001), cruciforms, hemi-catenated DNA, three and four way junctions,

telomeric t-loops and hairpins (Gohler et al., 2002; Stros et al., 2004). Heteroduplex joint intermediates that form during recombination have also been identified as structure-specific DNA binding sites (Janz et al., 2002) as p53 can interact with transfer regions during recombination and exonucleolytically correct heteroduplex intermediates following strand invasion. Another type of DNA that p53 can recognise is ssDNA, and p53 is capable of catalysing DNA renaturation and DNA strand transfer independent of ATP (Bakalkin et al., 1995; Bakalkin et al., 1994). As ssDNA can also be a product of DNA damage, p53 may recognise these sites as potentially dangerous lesions, so the direct interaction may lead to initiation of the appropriate response. Routine maintenance of genome stability by p53 in non-stressed cells also requires interactions with genomic DNA, particularly at sites prone to structural rearrangements such as sites of active metabolic processes including recombination intermediates and telomeric t-loops (Kim and Deppert, 2006).

As with sequence-specific binding, an intact tetramerisation domain is required for binding to unusual structures. DNA junctions that form in many of these structures are believed to be important elements required for this structure-dependent binding (Gohler et al., 2005). It has been proposed that this non-sequence specific binding by p53 may allow recruitment of cellular repair factors to sites of non-B DNA to remove unusual structures that may reduce genomic stability (Walter et al., 2005).

As mentioned previously, p53 is thought to be involved in the routine maintenance of sites in the genome that are prone to structural rearrangements. As described in Section 1.3.2, genomes contain regions of repeat sequences that are intrinsically unstable and can conform to unusual structures. Recent studies (Walter et al., 2005) investigated the ability of p53 to interact with structures formed by CTG•CAG (referring to sequences of both strands in a 5' – 3' direction) tracts that represent genomic DNA sequences involved in TRS related disorders. Interest in the mechanisms resulting in genome instability of CTG•CAG tracts is considerable as they are the cause of several human hereditary neurological diseases. It was discovered that repeat tracts comprised novel p53 binding sites and that interactions were primarily determined by secondary hairpin structures. The potential physiological implications are that TRS may represent non-canonical p53 binding sites and interactions may lead to activation of p53 regulated pathways. The downstream consequences of binding may be alterations in transcriptional regulation of target genes or, alternatively, binding may allow p53 to participate directly in processes such as DNA repair. The proposed idea of a direct participation in DNA repair is strengthened by correlations in binding to CTG•CAG tracts between p53 and the DNA MMR repair protein hMSH2 (Pearson et al., 1997). The detected

interaction between p53 and CTG●CAG containing hairpins did not require activation of p53, therefore maintenance of such genomic stability may be a function of latent p53.

Two models have been proposed by which this non-sequence specific binding of p53 may alter cellular events. The first model states that binding of p53 to unusual structures may transactivate downstream genes, usually resulting in arrest of the cell cycle or apoptosis. In this model, binding by the CTD would stimulate sequence-specific binding by the CDBD to specific sequences in the genome, meaning p53 would either need to be released from the unusual structure (Stenger et al., 1994), or require a process such as DNA looping in order to interact with distant sites. The second model proposes that various species of p53 molecules may exist within a cell due to different post-translational modifications and that certain p53 molecules may have more affinity towards unusual DNA structures than other species. Following binding to unusual structures, these p53 molecules may stimulate sequence-specific binding by other p53 molecules or they may interact with other nuclear proteins to produce the required response (Somasundaram, 2000).

1.6 Aims of this Study.

Unusual structural elements within DNA and damage-induced lesions can strongly influence various cellular processes and can ultimately compromise genomic stability. The overall aim of this project was to characterise macromolecular interactions that are essential to mechanisms believed to detect discontinuities in DNA. The recent observation that p53 interacts with CTG●CAG tracts that conform to secondary hairpin structures (Walter et al., 2005) is of considerable interest as this may demonstrate a direct role for p53 in the regulation of unstable genomic regions. One aim of this study was to investigate p53 DNA binding activity to DNAs that varied in sequence and structure, including oligonucleotides containing internal CTG repeat hairpins. Previously, *in vitro* assays have provided considerable information regarding p53-DNA interactions, however, significant doubt has been cast over the reliability of assays that require antibodies or mutated p53 molecules to detect specific binding. Hence, alternative *in vitro* precipitation DNA binding assays were employed to allow determination of p53-DNA interactions to DNAs exhibiting features that may be recognised as dangerous. It was also of interest to compare similar interactions in an *in vivo* environment and so a chromatin immunoprecipitation assay investigating p53-DNA binding was developed.

Recently, 3 Ku homologues in *S. coelicolor* were identified but as yet, very little is known regarding their physiological activity. However, it is anticipated that they constitute integral

roles during NHEJ in *S. coelicolor* by recognising and recruiting repair factors to DSBs. The DNA-binding activities of recombinant ScKu proteins were investigated using *in vitro* DNA binding assays alongside a previously characterised prokaryotic Ku from *M. tuberculosis* (Weller et al., 2002) to identify similarities and differences between these Ku proteins.

It was also of interest to see if NHEJ of *S. coelicolor* shows agreement with the current prokaryotic NHEJ viewpoint of homodimeric Ku participating in a two-component system. Structural studies would ascertain if these proteins existed as multiple protein species, and the ability of Ku to interact with each other and other NHEJ components was to be investigated using co-immunoprecipitation assays. An additional aim was to clone recombinant *scku* with promoter DNA into *S. coelicolor* so that *in vivo* analysis of protein expression and macromolecular interactions could be investigated.

CHAPTER 2

Materials and Methods

This chapter contains details of the experimental procedures used during the entirety of this study. This includes information regarding bacterial strains and plasmids used, procedures used to create and analyse specific plasmid and oligonucleotide DNAs, protein production and analysis and the biochemical techniques utilised to characterise protein functions. Constituents of media and buffers used are detailed in Section 2.15.

2.1 Growth and Maintenance of Bacterial Strains.

2.1.1 Bacterial Strains.

This study required a range of *E. coli* strains specialised for different functions including supercompetent cells for transformation of ligation products and expression strains used during production of recombinant proteins (Table 2.1). Selected strains of *S. coelicolor* M145 were also used during protein expression analysis (Table 2.2).

Table 2.1 Strains of *E. coli* used in studies.

<i>E. coli</i> strain	Genotype	Reference
BL21 (DE3) pLysS	F ⁻ <i>dcm ompT hsdS</i> (r _B ⁻ m _B ⁻) <i>gal λ</i> (DE3).	Novagen [®]
DH5α	Δ <i>lacU169 endA1 gyrA96 relA1 hsdR17 thi-1 supE44</i> .	(Sambrook et al., 1989)
CH3-Blue	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsd RMS-mcrBC</i>) φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 endA1 ara Δ139 Δ(ara-leu)7697 galU galK λ-rpsL</i> (Str ^R) <i>nupG</i> .	Bioline [©]
α-GOLD	F ⁻ <i>deoR endA1 recA1 relA1 gyrA96 hsdR17</i> (r _k ⁻ , m _k ⁺) <i>supE44 thi-1 phoA Δ(lacZYA argF)U169 Φ80lacZΔM15 λ</i> .	Bioline [©]
TOP10	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsd RMS-mcrBC</i>) φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 deoR araD139 Δ(ara-leu)7697 galU galK rpsL</i> (Str ^R) <i>endA1 nupG</i>	Invitrogen [™]
BW25113 / pIJ790	pIJ790 [<i>oriR101</i>] [<i>repA101(ts)</i>] <i>araBp-gam-be-exo</i> . Chromosome: (Δ(<i>araD-araB</i>)567 Δ <i>lacZ4787(::rrnB-4) lacIP-4000(lacI^ρ) λ⁻ rpoS369(Am) rph-1 Δ(rhaD-rhaB)568 hsdR514</i> .	(Datsenko and Wanner, 2000)

Table 2.2 Strains of *S. coelicolor* used in studies.

<i>S. coelicolor</i>	Genotype	Reference
M145	SCP1 ⁻ , SCP2 ⁻	(Hopwood et al., 1985)
M145 (Δ <i>scku1</i>)	SCP1 ⁻ , SCP2 ⁻ , <i>SCO5309</i> ⁻ (<i>scku1</i> ⁻)	(Handley, Bowater & Hutchings, Unpublished data)
M145 (Δ <i>scku2</i>)	SCP1 ⁻ , SCP2 ⁻ , <i>SCO0601</i> ⁻ (<i>scku2</i> ⁻)	(Handley, Bowater & Hutchings, Unpublished data)
M145 (Δ <i>scku1</i> / Δ <i>scku2</i>)	SCP1 ⁻ , SCP2 ⁻ , <i>scku1</i> ⁻ , <i>scku2</i> ⁻	(Handley, Bowater & Hutchings, Unpublished data)

2.1.2 Growth of Bacterial Cultures.

All media constituents are described in Table 2.9. *E. coli* were grown on inverted 1.5% Luria Bertani (LB) agar plates or in liquid LB shaking at 200 rpm unless otherwise stated. Cultures were grown at 37 °C except during protein over-expression (25 °C) or growth of *E. coli* BW25113 / pIJ790 prior to recombination of PCR resistance cassettes (30 °C). *S. coelicolor* strains were grown at 30 °C on inverted Mannitol soya flour medium (SFM) or LB agar solid media or in liquid Tryptone soya broth (TSB) at 200 rpm.

2.2 DNA Procedures.

2.2.1 Transformation of Plasmid DNA into *E. coli*.

Constituents of all buffers can be found in Tables 2.10 – 2.13. Typically, plasmid DNA was added to 25 μ l *E. coli* strains (Table 2.1) and incubated on ice for 30 minutes. Bacterial cells were then heat shocked at 42 °C for 2 minutes and placed on ice for 2 minutes. SOC media (300 μ l) was added to the cells, which were then incubated at 37 °C for 1 hour at 200 rpm. Transformed cells (100 μ l) were then plated onto selective LB agar plates and incubated overnight at 37 °C.

2.2.2 Bacterial Plasmids and Cosmids used During Studies.

Table 2.3 Plasmids and Cosmids used During Studies.

Plasmid name	Description	Source
pET28(a)	An expression vector containing an N-terminal His-Tag [®] /T7-Tag [®] (6 His residues) and an optional C-terminal His-Tag [®] (6 His residues). It contains T7 forward and reverse primer sites for sequencing or PCR screening. P _{T7lac} , Kan ^R .	Novagen [®]
pRB16(b)	Expression vector containing an N-terminal His-Tag [®] (10 His residues). It contains T7 forward and reverse primer sites for sequencing or PCR screening. P _{T7lac} , Amp ^R .	Novagen [®]
pRB209	<i>SCO5301</i> (<i>scku1</i>) gene cloned into pET28(a).	(Sayer, 2008)
pRB210	<i>SCO0601</i> (<i>scku2</i>) gene cloned into pET28(a).	(Sayer, 2008)
pRB211	<i>SCP1.285c</i> (<i>scku3</i>) gene cloned into pET28(a).	(Sayer, 2008)
pRB219	<i>SCO7345</i> (<i>scligD</i>) gene cloned into pTrc99a.	(Sayer, 2008)
pRB141	<i>mtku</i> gene cloned into pET-16(b).	(Weller et al., 2002)
pUC18	Vector containing pMB1 replicon <i>rep</i> , <i>bla</i> gene (ampicillin resistance) and <i>E. coli lac</i> operon region consisting of CAP protein binding site, Plac promoter, <i>lac</i> repressor binding site and 5'-terminal of <i>lacZ</i> gene.	(Yanisch-Perron et al., 1985)
pL1	pUC19 containing recognition sequence for BbvCI restriction enzyme.	(Gowers et al., 2005)
pSET152	ΦC31-derived integration vector that targets an <i>attB</i> site in <i>S. coelicolor</i> <i>SCO3798</i> gene. Apra ^R .	(Combes et al., 2002)
pMS82	ΦBT1-derived integration vector that targets an <i>attB</i> site in <i>S. coelicolor</i> <i>SCO4848</i> gene. Hyg ^R .	(Gregory et al., 2003)
pCR [®] -Blunt II-TOPO [®]	A pUC derived plasmid used for cloning of blunt-ended DNA featuring M13 forward and reverse primer sites for sequencing or PCR screening. P _{lacZα} , Kan ^R , Zeo ^R .	Invitrogen [™]
St6G9	WT cosmid containing <i>scku1</i> .	(Redenbach et al., 1996)
St5G5	WT cosmid containing <i>scku2</i> .	(Redenbach et al., 1996)
pRB240	pCR [®] -Blunt II-TOPO [®] with uncharacterised modification (circularised empty vector that does not cause lethality in <i>E. coli</i>).	This study

The cosmids used during this study were from the Supercos-1 library that contains chromosomal DNA of a plasmid-free derivative of *S. coelicolor* A3(2) (Redenbach et al., 1996).

2.2.3 Plasmid Extraction from *E. coli*.

Bacterial cultures containing relevant plasmid DNA were grown overnight in 10 ml LB media containing selective antibiotics. Plasmids were extracted using the Promega (Madison WI, USA) Wizard SV+ mini-prep kit and according to the manufacturer's protocol. Plasmids were eluted with 1 x TE buffer and stored at -20 °C.

2.2.4 Clean-up of Extracted Plasmids.

For all experiments involving plasmid DNA and ScKu proteins, plasmids were required to be purified following extraction or restriction digestion to prevent significant degradation of ScKu proteins. QIAquick[®] PCR purification Kit (Qiagen) was used according to the manufacturer's spin protocol and plasmids were eluted in 1 x TE buffer and stored at -20 °C.

2.2.5 Polymerase Chain Reaction (PCR).

2.2.5.1 Amplification of Promoter Regions Following Chromatin Immunoprecipitation.

PCR was performed for detection of promoter regions during Chromatin Immunoprecipitation (ChIP) assays (Section 2.13). For all reactions, 22.5 µl of 1.1 x Pre-Aliquoted ReddyMix[™] PCR Master Mix (ABgene[®] UK) was used. In a final volume of 25 µl this mix contained: 0.625 units Thermoprime Plus DNA Polymerase, 75 mM Tris-HCl (pH 8.8 at 25 °C), 20 mM (NH₄)₂SO₄, 1.5 mM MgCl₂, 0.01% (v/v) Tween[®] 20 and 0.2 mM each of dATP, dCTP, dGTP and dTTP. To the mix, 2% Dimethyl sulfoxide (DMSO), 300 nM of both forward and reverse primers and varying volumes of template DNA were added. All reactions were performed in a Mastercycler[®] Personal PCR machine (Eppendorf[®]).

Table 2.4 PCR primers used to amplify genomic DNA from ChIP assays.

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>p21</i>	GTGGCTCTGATTGGCTTTCTG	CTGAAAACAGGCAGCCCAAG
<i>Mdm2</i>	GGTTGACTCAGCTTTTCCTCTTG	GGAAAATGCATGGTTTAAATAGCC
<i>GAPDH</i>	GTATTCCCCCAGGTTTACAT	TTCTGTCTTCCACTCACTCC

For the amplification of p21 and Mdm2 promoter regions, the thermal cycler program was as follows: 30 seconds at 98 °C for initial denaturation, then 40 cycles of: [20 second denaturation at 94 °C, 20 second annealing at 60 °C and 1 minute elongation at 68 °C]. Following this cycle, a 10 minute final extension was performed at 68 °C.

For the amplification of GAPDH promoter region, the thermal cycler program was as follows: 30 seconds at 98 °C for initial denaturation, then 40 cycles of: [20 second denaturation at 94 °C, 20 second annealing at 55 °C and 1 minute elongation at 68 °C]. Following this cycle, a 10 minute final extension was performed at 68 °C.

2.2.5.2 Amplification of DNA for Cloning.

Preliminary PCRs were performed using GoTaq® Flexi DNA Polymerase (Promega) to ascertain optimum conditions for subsequent PCRs, which were performed using Expand High Fidelity PCR System (Roche). Reaction mixes typically contained: 1.25 U GoTaq® DNA Polymerase, 1 x GoTaq® Flexi buffer, 1 mM MgCl₂, 0 – 12% DMSO, 300 nM of both forward and reverse primers (Table 2.5), varying concentrations of template DNA (cosmid (St6G9 for *sku1* or St5G5 for *sku2*) or *S. coelicolor* genomic DNA) and 0.2 mM each of dATP, dCTP, dGTP and dTTP.

The GoTaq® Flexi PCR cycles were as follows: 3 minutes at 95 °C for initial denaturation, then 30 cycles of: [30 second denaturation at 95 °C, 30 second annealing at 59 °C and 2 minute elongation at 72 °C]. Following this cycle, a 7 minute final extension was performed at 72 °C.

Table 2.5 PCR primers used to amplify *SCO5309* and *SCO0601* genes. Black text indicates regions of complementarity towards template DNA, blue text indicates restriction sites, red text indicates stop codon and green text indicates polyhistidine tag sequence.

Gene	Target Vector	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>SCO5309</i>	pMS82	ACTAGTATACCGCAGGAA GGAGACCGCGC	GATATCTCAGTGATGGTGATGGTGG TGGGCGCTGCGCCTGCGGG
<i>SCO5309</i>	pSET152	TCTAGAATACCGCAGGAA GGAGACCGCGC	GATATCTCAGTGATGGTGATGGTGG TGGGCGCTGCGCCTGCGGG
<i>SCO0601</i>	pMS82	ACTAGTCAGGGGTGAGGC GGCCGG	GATATCTCAGTGATGGTGATGGTGG TGGGCCGCGGTCGTCGTCTTCTTC
<i>SCO0601</i>	pSET152	TCTAGACAGGGGTGAGGC GGCCGG	GATATCTCAGTGATGGTGATGGTGG TGGGCCGCGGTCGTCGTCTTCTTC

DNA to be cloned into integration vectors was amplified using the Expand High Fidelity PCR System manufactured by Roche, using the Mastercycler® Personal PCR machine (Eppendorf). Reactions contained 2.6 U Expand High Fidelity Enzyme Mix, 1 x Expand High Fidelity Buffer (including MgCl₂), 300 nM of both forward and reverse primers (Table 2.5), 20 ng template cosmid DNA (St6G9 for *scku1* or St5G5 for *scku2*) and 0.2 mM each of dATP, dCTP, dGTP and dTTP. The final volume was made up to 50 µl with dH₂O.

Expand High Fidelity PCR cycles were as follows: 3 minutes at 95 °C for initial denaturation; 10 cycles of: [94 °C, 30 second denaturation; 59 °C, 30 second annealing; 72 °C, 2 minute elongation]; 20 cycles of: [94 °C, 30 second denaturation; 59 °C, 30 second annealing; 72 °C, 2 minute, 5 seconds elongation with an accumulative 5 seconds on each cycle]; 72 °C, 10 minute final extension.

2.2.5.3 Colony PCR.

PCR was used to screen transformants containing the required cloned DNA. Single colonies were picked with a sterile pipette tip that was used to thoroughly mix individual Pre-Aliquoted ReddyMix™ PCR Mix (ABgene) reaction tubes containing 300 nM of forward and reverse primers. Reactions were performed using the same cycle parameters as described in Section 2.2.5.2 during GoTaq® Flexi PCR.

2.2.6 Agarose Gel Electrophoresis.

Plasmid DNA was visualised on 0.8 or 1% agarose gels in 100 ml of 1 x TAE buffer containing 0.5 µg / ml ethidium bromide. An appropriate volume of 6 x loading dye was added to DNA and electrophoresis was performed at 120 V. Gels were visualized with UV light using a transilluminator (BioRAD, Munich, Germany). If additional staining was required, gels were placed in 200 ml water containing 0.5 µg / ml ethidium bromide for 20 minutes and washed for 30 minutes in dH₂O before visualisation.

2.2.7 DNA Purification Following Agarose Gel Electrophoresis.

Agarose gels were visualised on a transilluminator and sections of gels containing DNA fragments were excised using a scalpel, with minimum exposure to UV light. DNA was then purified from gel sections using QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's protocol. Typically, DNA was eluted in 30 µl of elution buffer.

2.2.8 Cloning of *scku* Genes.

2.2.8.1 Cloning *scku* Genes into pCR®-Blunt II-TOPO®.

Zero Blunt® TOPO® PCR Cloning (Invitrogen) was selected for cloning reactions as this strategy is highly efficient for the direct insertion of blunt-end PCR products into vector DNA. These vectors would act as holding vectors prior to cloning of *scku* genes into integration vectors. The TOPO® cloning reaction relies on the activity of topoisomerase I from *Vaccinia* virus that attaches to duplex DNA and cleaves the phosphodiester backbone, forming a covalent bond between the 3' phosphate of the DNA and a tyrosyl residue of the topoisomerase. This reaction can be subsequently reversed and the DNA sealed via the 5' hydroxyl of the cleaved strand attacking the phospho-tyrosyl bond (Shuman, 1994). The pCR®-Blunt II-TOPO® vector plasmid is supplied linearised with *Vaccinia* virus topoisomerase I covalently bound to the 3' ends of the DNA. Additionally, pCR®-Blunt II-TOPO® encodes the lethal *E. coli* gene *ccdB* that becomes disrupted following ligation of a PCR product into the plasmid and prevents its expression, allowing for accurate screening of successful recombinants.

Cloning of amplified *scku* genes into pCR®-Blunt II-TOPO® was performed according to the manufacturer's protocol: 4 µl of fresh PCR product (120 ng) was added to 1 µl of 1.2 M NaCl, 0.06 M MgCl₂ and 1 µl of TOPO® vector and the reaction mix was incubated at 22 °C for 10 minutes. 2 µl of the TOPO® cloning reaction was then added to a vial of *E. coli* TOP10 cells and transformed as detailed in Section 2.2.1.

2.2.8.2 DNA Sequencing.

Approximately 200 ng of pCR®-Blunt II-TOPO® construct DNA was sequenced at Genome Enterprise Limited, Norwich, UK, using primers specific for SP6 and T7 promoter sites that can produce sequence fragments of around 1000 bases. Sequence analysis was performed using DNASTar 6 (Lasergene).

2.2.8.3 Cloning *scku* Genes into Integration Vectors.

Phage-derived vectors pMS82 and pSET152 were to be used for direct integration of *scku* genes into the *S. coelicolor* chromosome. For subcloning of *scku* genes that were situated in TOPO® vectors into integration vectors, both sets of vectors required 2 digestions with relevant restriction enzymes (Table 2.6). Following primary digests, products were cleaned

(Section 2.2.4), a small sample analysed by electrophoresis (Section 2.2.6) to confirm digestion, and then remaining products were subjected to a second digestion. Restriction digest products were subject to electrophoresis and the required DNA fragments were excised and purified from agarose gels (Section 2.2.7).

Table 2.6 Restriction enzymes used to clone *scku* genes into integration vectors.

Gene	Integration Vector	Restriction Enzymes
<i>scku1</i>	pMS82	<i>Spe1</i> and <i>EcoRV</i>
<i>scku1</i>	pSET152	<i>Xba1</i> and <i>EcoRV</i>
<i>scku2</i>	pMS82	<i>Spe1</i> and <i>EcoRV</i>
<i>scku2</i>	pSET152	<i>Xba1</i> and <i>EcoRV</i>

Inserts and integration vectors were then combined at a molar ratio of 1:3 (vector: insert) and ligation reactions were performed at 20 °C for 16 hours that contained: 1 U of T4 DNA ligase and 1 X T4 DNA ligase buffer (Roche) in 10 µl reaction volumes that also contained relevant DNAs. 5 µl of the reaction mix was transformed into 50 µl of CH3-Blue cells (Section 2.2.1) and successful constructs were confirmed by colony PCR (Section 2.2.5.3) and restriction digest analysis.




2.3 Oligonucleotide Preparation and Analysis.

2.3.1 Preparing DNA Substrates for *in vitro* DNA Binding Assays.

Various oligonucleotide substrates were employed during *in vitro* DNA binding assays. Single-stranded DNA (ssDNA) oligonucleotides were purchased from Eurogentec, resuspended in 1 x TE buffer and stored at -20 °C. Double-stranded DNA (dsDNA) oligonucleotides, consisting of 1 biotinylated and 1 unbiotinylated strand, were produced by combining specific ssDNA oligonucleotides in a 1:4 ratio (biotinylated: unbiotinylated). Although excess unbiotinylated ssDNA was present in samples following annealing, wash procedures during assays would insure these DNAs would not interfere with experiments. An exception to this was a dsDNA molecule (dsDNA+ biotin) that consisted of 2 biotinylated strands of which both strands were added in equal amounts during annealing. Prior to annealing, KCl was added to oligonucleotides to a final concentration of 50 mM. For annealing, oligonucleotides were placed in a water bath at 100 °C for 8 minutes. Oligonucleotides were then allowed to cool to 20 °C over 6 hours and stored at -20 °C.

Table 2.7 Oligonucleotide sequence and structural information. Green text refers to biotinylated DNA strands, black text represents unbiotinylated DNA strands and red text indicates unusual structural features. Blue circles on DNA termini represent biotin molecules and orange circles represent conjugated fluorescein fluorophore molecules. All DNA molecules were produced as described in this section using optimal conditions and analysed as described in Section 2.3.2.

Oligonucleotide	Structure
<p>p53 p21-RE</p> <p>5' B-caacGGCCATCAGGAACATGTCCCAACATGTTGAGCTCTGG- 3'</p> <p>3' - CCGGTAGTCCTTGTACAGGGTTGTACAACCTCGAGACC- 5'</p>	
<p>p53 NBS1</p> <p>5' B-caacGCTGCCTGTTTTCAGGGAGGAAGGGGATGGTAGGAGA- 3'</p> <p>3' - CGACGGACAAAAGTCCCTCCTTCCCCTACCATCCTCT- 5'</p>	
<p>p53 CTG dsDNA</p> <p>5' B-caacCCGCGGTACCATCAGCAGCAGCAGCAGCAGCAGTACCTAAGGCCCC- 3'</p> <p>3' - GGCGCCATGGTAGTCGTCGTCGTCGTCGTCATGGATTCCGGGG- 5'</p>	
<p>p53 (CTG)₁ hairpin (HP)</p> <p>5' B-caacCCGCGGTACCAT===TACCTAAGGCCCC- 3'</p> <p>3' - GGCGCCATGGTAGTCATGGATTCCGGGG- 5'</p>	
<p>p53 (CTG)₃ hairpin (HP)</p> <p>5' B-caacCCGCGGTACCAT=====TACCTAAGGCCCC- 3'</p> <p>3' - GGCGCCATGGTAGTCGTCGTCGTCGTCGTCATGGATTCCGGGG- 5'</p>	
<p>p53 (CTG)₇ hairpin (HP)</p> <p>5' B-caacCCGCGGTACCAT=====TACCTAAGGCCCC- 3'</p> <p>3' - GGCGCCATGGTAGTCGTCGTCGTCGTCGTCGTCATGGATTCCGGGG- 5'</p>	
<p>p53 (GAA)₇ hairpin (HP)</p> <p>5' B-caacCCGCGGTACCAT=====TACCTAAGGCCCC- 3'</p> <p>3' - GGCGCCATGGTAAGAAGAAGAAGAAGAAGATGGATTCCGGGG- 5'</p>	
<p>p53 (CGG)₇ hairpin (HP)</p> <p>5' B-caacCCGCGGTACCAT=====TACCTAAGGCCCC- 3'</p> <p>3' - GGCGCCATGGTAGGCGGGCGGCGGCGGCGGCGCATGGATTCCGGGG- 5'</p>	
<p>p53 (TAT)₇ hairpin (HP)</p> <p>5' B-caacCCGCGGTACCAT=====TACCTAAGGCCCC- 3'</p> <p>3' - GGCGCCATGGTATATTATTATTATTATTATATGGATTCCGGGG- 5'</p>	

dsDNA1 5' B- caacTTTGAGGTGCGTGTGTTGTGCCTGTCCTGGG- 3' 3' - AAATCCACGCACAAACACGGACAGGACCC- 5'	
dsDNA1+ G/T mm 5' B- caacTTTGAGGTGCGTGTGTTGTGCCTGTCCTGGG- 3' 3' - AAATCCACGTACAAACACGGACAGGACCC- 5'	
dsDNA1+ G/A mm 5' B- caacTTTGAGGTGCGTGTGTTGTGCCTGTCCTGGG- 3' 3' - AAATCCACGACAAACACGGACAGGACCC- 5'	
dsDNA1+ 1 base 5' B- caacTTTGAGGTGC=GTGTTTGTGCCTGTCCTGGG- 3' 3' - AAATCCACGTACAAACACGGACAGGACCC- 5'	
dsDNA1+ loop 5' B- caacTTTGAGGTGC=====GTGTTTGTGCCTGTCCTGGG- 3' 3' - AAATCCACGTGTGTGTGCACAAACACGGACAGGACCC- 5'	
dsDNA1+ bubble 5' B- caacTTTGAGGTGCGTGTGTTGTGCCTGTCCTGGG- 3' 3' - AAATCCACGTATGGTCTGGACAGGACCC- 5'	
dsDNA1+ biotin 5' B- caacTTTGAGGTGCGTGTGTTGTGCCTGTCCTGGG- 3' 3' - AAATCCACGCACAACTCGGACAGGACCC-B 5'	
dsDNA+ fluorophore 5' B- caacACTTGAATCGAGCGA- 3' 3' - TGAACCTAGCTCGCT-FL 5'	
p53 ssDNA 5' B- caacCCGCGGTACCATTACCTAAGGCCCC- 3'	
ssDNA1 5' B- caacTTTGAGGTGCGTGTGTTGTGCCTGTCCTGGG- 3'	
ssDNA+fluorophore 5' - FL-CCCAGGACAGGCACAAACACGCACCTCAAA- 3'	

2.3.2 Analysis of Oligonucleotides by Native and Denaturing Gel Electrophoresis.

To determine the success of oligonucleotide annealing, including incorporation of unusual structural features, all oligonucleotide substrates (typically 50 fmol) were analysed using native and denaturing gel electrophoresis. Migration of oligonucleotides during electrophoresis is determined not only by size and nucleic acid composition, but also by

structure. Therefore, this was a suitable method for determining if secondary structures had been incorporated during the annealing process (Chastain et al., 1995). Native and denaturing polyacrylamide gels (10% and 12%) were produced as described (Sambrook et al., 1989). For native electrophoresis, an appropriate volume of native loading buffer was added to samples and electrophoresis was performed in Bio-Rad Mini-PROTEAN® 3 cells at 50 V for 3 hours using 1 x TBE as running buffer. For denaturing gels, an appropriate volume of 1 x stop solution was added to samples which were then heated at 100 °C for 5 minutes. Electrophoresis was performed in Bio-Rad Mini-PROTEAN® 3 cells at 400 V for 30 minutes using 1 x TBE as running buffer.

2.3.3 Transfer to Membrane and Chemiluminescent Detection of Oligonucleotides.

Oligonucleotides were transferred onto 0.2 µm nylon membrane (Pierce) using a semi-dry blotting system (Scie-Plas Ltd) according to the manufacturer's instructions. Membrane and blotting paper were soaked in 1 x TBE prior to transfer. Transfers were performed at 400 mA for 40 minutes for each 10 x 8 cm gel. Following transfer, oligonucleotides were cross-linked to the membrane using UV Stratalinker® 1800 (Stratagene®) using auto cross-link function. Chemiluminescent detection was performed using the Chemiluminescent Nucleic Acid Detection Module (Pierce) and according to the manufacturer's protocol. Luminescence was detected either using chemiluminescence detection program on a Universal Hood II (Bio-Rad) or by developing Medical X-ray film (Fuji) using Compact X4 (Xograph Imaging Systems) developer.

2.4 Protein Over-expression and Purification.

2.4.1 pET Expression System.

The pET expression system was chosen for protein over-expression as expression can be directly controlled and a high yield of protein can be obtained. *E. coli* BL21 (DE3) contains genetically engineered sites in the chromosome required to control protein expression. A T7 RNA polymerase gene that has been engineered onto the chromosome is used to transcribe the gene of interest. Expression of the T7 RNA polymerase is regulated by both the *lacI* gene, whose product inhibits transcription of the T7 RNA polymerase gene when bound to the *lac* promoter, and also a T7 lysozyme that prevents activity at low levels (Moffatt and Studier, 1987). The gene to be expressed is cloned into a plasmid downstream of a T7 promoter and this plasmid can then be transformed into *E. coli* BL21 (DE3). Expression of the gene of interest is greatly increased following addition of isopropyl-β-D-thiogalactopyranoside (IPTG), which prevents LacI binding to the *lac* promoter and thereby allowing continual

expression of T7 RNA polymerase, which then transcribes the gene of interest. The T7 promoter is a viral promoter and it transcribes rapidly in the presence of T7 RNA polymerase, thereby making large quantities of the desired protein quickly.

2.4.2 Protein Over-expression.

Single *E. coli* BL21 pLysS (DE3) colonies containing the plasmid of interest were inoculated into 10 ml selective LB media and grown overnight at 37 °C with shaking at 200 rpm. Resulting cultures were used to inoculate 1 L of selective LB media which were then grown at 25 °C with shaking at 200 rpm. Once growth had reached mid logarithmic growth phase ($OD_{600} = 0.4 - 0.6$), IPTG was added at a final concentration of 0.4 mM to induce recombinant protein over-expression. Cultures were then incubated overnight at 25 °C with shaking at 200 rpm. Cell pellets were harvested by centrifugation at 6000 rpm (Beckman Coulter™ Avanti™ centrifuge, JLA 8.1000 rotor) for 20 minutes at 4 °C. Supernatant was removed and the pellet was resuspended in 100 ml of this supernatant and split into 25 ml aliquots and centrifuged at 5000 rpm (Beckman Coulter™ Allegra™ 25R centrifuge with TS-5.1-500 swing bucket rotor) for 20 minutes at 4 °C. Supernatant was removed and discarded and pellets stored at -80 °C.

2.4.3 Protein Purification.

2.4.3.1 Protein Extraction.

Cell pellets were defrosted from storage at -80 °C and resuspended in 20 ml of 1 x binding buffer for every 1 L pellet. Resuspended cells were then divided and decanted into 2 x 50 ml centrifuge tubes and sonicated 10 x 10 seconds with 1 minute rest on ice between each sonication. Cells were centrifuged at 16000 rpm (Beckman Coulter™ Avanti™ centrifuge, JA 25.50 rotor) for 40 minutes at 4 °C and the supernatant was removed for purification of recombinant protein. A small sample of pellet was removed and stored at -20 °C for analysis.

2.4.3.2 Novagen His·Bind® Column Purification.

Recombinant Ku and ligase proteins in this study contained N-terminal polyhistidine tags so purification could be achieved using His·Bind® columns (Novagen®). All steps were performed at 4 °C following manufacturer's instructions. Initially, columns were washed with 10 ml 1 x binding buffer then 10 ml of 1 x charge buffer was added to each column and left to run through. To equilibrate columns, 5 ml of 1 x binding buffer was added. Cell supernatant (100 µl) (Section 2.4.3.1) was stored as sample S1 for later analysis, the

remaining supernatant was filtered through a 0.45 µm filter, of which 100 µl was stored as sample S2. The remaining supernatant was loaded onto equilibrated columns and flow-through was collected as sample S3. The column was washed with 10 ml of 1 x binding buffer and the flow-through was collected (sample S4), followed by 10 ml of 1 x wash buffer (flow-through kept as sample S5). Protein was eluted from columns using 1 x elution buffer in 8 x 1 ml fractions (samples F1-F8) and 2 x 2 ml fractions (F9-10). Fractions were analysed (Section 2.6) to discern which contained the most protein.

2.4.3.3 GE Healthcare HisTrap™ HP Purification.

HisTrap™ HP cartridges (GE Healthcare) were used in conjunction with Äktaprime™ plus (GE Healthcare). Initially, HisTrap™ HP cartridges were prepared by washing with 15 ml dH₂O followed by charging with 10 ml 1 x charge buffer. Cartridges were again washed with 15 ml dH₂O and inserted into Äktaprime™ plus. Äktaprime™ A line was then washed with 20 ml 1 x binding buffer (10 ml / min), followed by a cartridge wash with 50 ml 1 x binding buffer (5 ml / min). Supernatant from protein extraction (Section 2.4.3.1) was filtered through a 0.45 µm filter and added to the cartridge at 1 ml / min, followed by a 10 ml 1 x binding buffer wash at 1 ml / min. Äktaprime™ plus B line was then washed with 1 x elution buffer followed by A line wash with 20 ml 1 x wash buffer (10 ml / min). Prior to protein elution, the cartridge was washed with 5 ml 1 x wash buffer (1 ml / min) and the fraction collected. A step gradient was used to elute recombinant proteins, beginning with 100% 1 x wash buffer, but increasing by 20% 1 x elution buffer every 5 ml. Elution was performed at 1 ml / min for 30 minutes and 1 ml fractions were collected. Äktaprime™ plus system was washed with dH₂O and cartridges were stripped with 10 ml 1 x strip buffer and stored in 20% ethanol.

2.4.3.4 Desalting and Buffer Exchange.

PD10 columns (Amersham Biosciences) were used for buffer exchange and desalting of protein fractions (2.5 ml) containing the most protein. The manufacturer's protocol was then followed. Eluted protein was either snap frozen in liquid nitrogen or glycerol was added to a final concentration of 20% before storage at -80 °C. PD10 columns were washed with 20% ethanol and stored at 4 °C for re-use.

2.4.3.5 Fast Performance Liquid Chromatography (FPLC).

Gel filtration chromatography was occasionally used to further purify recombinant proteins. A Liquid Chromatography Controller LCC-500 and HiLoad™ 16'60 Superdex™ 75 prep grade column (Pharmacia Biotech) were used to perform FPLC. Protein sample fractions from purification processes of volumes between 1 to 2 ml were loaded into the column and the protein was processed at 1 ml / min. Absorbance was recorded at 0.2 cm / min and protein samples were collected following filtration.

2.4.3.6 Homogenisation of High Five Insect Cells.

High Five insect cells (Invitrogen™) were supplied that had been infected with recombinant baculovirus controlling human wild-type p53 expression. Prior to purification, cells had been stored at -80 °C for several months. Purification was performed as previously described (Dudenhofer et al., 1998). Briefly, cells were washed in ice cold PBS followed by incubation in ice cold buffer A for 50 minutes allowing cells to swell. Swollen cells were broken up using a Dounce homogeniser and samples left on ice for a further 30 minutes. Cells were then centrifuged at 6500 rpm (Beckman Coulter™ Allegra™ 25R centrifuge with TS-5.1-500 swing bucket rotor) at 4 °C for 30 minutes in order to harvest nuclei. A small sample (sample A) was taken from the supernatant and the remaining pellet was resuspended in buffer B and stored on ice for 40 minutes. Cells were then centrifuged at 6500 rpm at 4 °C for 30 minutes and supernatant was collected (sample B). A small volume of buffer B2 was added to the remaining pellet and the sample was vortexed and left on ice for 30 minutes. The sample was again centrifuged at 6500 rpm at 4 °C for 20 minutes and supernatant collected (sample C). Glycerol was added to samples A, B and C to a final concentration of 20% and stored at -80 °C.

2.5 Bacterial Cell Extract Analysis.

2.5.1 *E. coli* and *S. coelicolor* Cell Extract Harvesting.

Media containing selective antibiotics were inoculated with bacterial colonies and grown for 16 or 40 hours. Cultures were centrifuged at 5000 rpm for 15 minutes at 4 °C (Beckman Coulter™ Allegra™ 25R centrifuge with TS-5.1-500 swing bucket rotor) and then resuspended in 1 x binding buffer containing 1 x protease inhibitors (Complete Mini, EDTA-free, Roche). Cells were sonicated 3 x 5 seconds, with 1 minute rest on ice between each

sonication, and then centrifuged at 5000 rpm for 10 minutes at 4 °C and supernatant was collected.

2.5.2 *S. coelicolor* Spore Harvesting, Viability Count and Extract Analysis.

For analysis of spore viability of the various *S. coelicolor* strains, single colonies were streaked on to 3 SFM plates so that the entire plate surface would be covered and were grown for 5 days at 30 °C. Glycerol (2 ml of 20%) was applied to each plate and spores were suspended in the glycerol with a cotton bud and the resultant spore solutions were collected. Spores were then filtered using cotton wool and centrifuged at 5000 rpm for 15 minutes at 4 °C. Following supernatant decanting, spores were resuspended in 1 ml of 20% glycerol.

Viability counts were attained by plating 100 µl of 10^{-6} and 10^{-12} dilutions on to non-selective LB plates and grown at 30 °C for 2 days. The colonies were counted using white light on a Universal Hood II (Bio-Rad) and viability was expressed as the number of colony forming units.

For spore extract harvesting, 700 µl of spore solutions were centrifuged at 5000 rpm for 10 minutes at 4 °C and then resuspended in 500 µl of 1 x binding buffer containing 1 x protease inhibitors. Solutions were decanted into 2 ml centrifuge tubes containing a small amount of <106 µm glass beads (Sigma-Aldrich®) and were homogenised using FastPrep-24® (MP Biomedicals) by 6 x 30 second cycles with 5 minutes on ice between cycles. Solutions were briefly centrifuged at 2000 rpm for 30 seconds and supernatant was removed and centrifuged at 13000 rpm for 20 minutes at 4 °C. Supernatants were collected and stored at -20 °C.

2.6 Protein Analysis.

2.6.1 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE).

Gels for protein analysis consisted of 10% or 12% polyacrylamide and were produced as described (Sambrook et al., 1989). Electrophoresis was performed in Bio-Rad Mini-PROTEAN® 3 cells at 150 V using 1 x Tris-glycine SDS as running buffer. Prior to electrophoresis, an appropriate volume of 4 x Tris-glycine SDS-PAGE loading buffer was added to samples, which were then heated at 100 °C for 8 minutes. Bio-Rad prestained SDS-PAGE Standard Low Range was used as a molecular weight marker.

2.6.2 Coomassie Staining of SDS-PAGE Gels.

SDS-PAGE gels were washed twice in dH₂O and then stained in either Coomassie blue or InstantBlue (Lucerna-Chem). For Coomassie blue development, gels were incubated for 2 hours with gentle agitation and once stained, gels were de-stained by twice boiling in 600 ml of dH₂O by heating in a microwave at maximum power for 10 minutes. For InstantBlue staining, gels were incubated for 20 minutes and washed briefly in dH₂O. Images of gels were taken using white light on a Universal Hood II (Bio-Rad).

2.6.3 Western Blotting.

Western blots were performed using a semi-dry blotting system (Scie-Plas Ltd) and according to manufacturer's instructions. Proteins were transferred from SDS-PAGE gels onto 0.2 µm pure nitrocellulose membrane (Bio-Rad) (pre-soaked in 1 x blot buffer) typically at 50 mA for 2 hours for each 9 cm x 6 cm gel. Immunodetection was performed using a ProtoBlot® II AP system (Promega) following the manufacturer's protocol. Western blue (Promega) was used to develop all nitrocellulose membrane western blots.

Table 2.8 Antibodies used during immunodetection of proteins.

Protein	Primary antibody	Secondary antibody(s)
p53	Mouse monoclonal p53 (DO-1) (1:100,000) (Santa Cruz).	Anti-Mouse - Alkaline Phosphatase (AP) conjugate (1:5000) (Promega) / Anti-Mouse - Horse Raddish Peroxidase (HRP) conjugate (1:1250) (GE healthcare).
Phosphorylated p53	Mouse monoclonal Phospho-p53 (Ser15) (1:1000) (Cell signalling).	Anti-Mouse – AP conjugate (1:5000) (Promega) / Anti-Mouse – HRP conjugate (1:1250) (GE healthcare).
β-actin	Mouse polyclonal β-actin (1:1000) (Cell signalling).	Anti-Mouse - HRP conjugate (1:1250) (GE healthcare).
ScKu1, ScKu2, ScKu3, MtKu	Anti-His tag (1:3000) (GE healthcare).	Anti-Mouse – AP conjugate (1:5000) (Promega).
ScKu1	ScKu1 rabbit polyclonal (1:1000 – 1:10000) (Cambridge Research Biochemicals).	Anti-Rabbit – AP conjugate (1:10000) (Sigma-Aldrich®).
ScKu2	ScKu2 rabbit polyclonal (1:1000 – 1:10000) (Cambridge Research Biochemicals).	Anti-Rabbit – AP conjugate (1:10000) (Sigma-Aldrich®).

2.6.4 Protein Concentration Determination using Bradford Assay.

Bradford assay (Bio-Rad) was used to determine protein concentrations. The manufacturer's protocol was followed with BSA used to create a standard curve. All A_{595} readings were recorded on a BioPhotometer (Eppendorf).

2.6.5 Calculating Protein Absorbance at 280 nm.

During Analytical Ultracentrifuge experiments (Section 2.11), protein absorbance at 280 nm was measured. U3310 spectrophotometer (Hitachi) and UV solutions application (Hitachi) were used to measure recombinant protein absorbance. Parameters used were as follows – start wavelength 340 nm, end wavelength 200 nm, scan speed 300 nm / min and path length 10 mm. Protein storage buffer (200 μ l) was measured as a baseline before 100 μ l of protein was added to the same 200 μ l buffer. Absorbance was recorded at 280 nm.

2.7 *In vitro* Streptavidin Magnetic Bead (SMB) DNA Binding Assay.

2.7.1 *In vitro* SMB DNA Binding Assay Principle and Mechanism.

To determine DNA binding properties of proteins investigated in this study, a new technique based on previous work (Palecek et al., 2004) was developed. The basic mechanism is demonstrated in Figure 2.1 and all reactions were performed at 30 °C. An appropriate amount of Streptavidin coated Magnetic Beads (SMBs) (Invitrogen™ Dynabeads® M-280) were added to 1.5 ml microcentrifuge tubes and were washed with 100 μ l 1 x SMB wash buffer. Microcentrifuge tubes were then positioned on a Neodymium rare earth permanent magnet Magnetic Separation Rack (New England BioLabs® Inc.) for 1 minute to allow SMBs to collect in the tube adjacent to the magnet. Supernatant was removed and discarded and this wash step was repeated. SMBs were then resuspended in 25 μ l of 1 x SMB wash buffer containing an appropriate amount of biotinylated oligonucleotides to be used as the protein substrate. Oligonucleotides were added in excess so that total oligonucleotide substrate available in each reaction was dependent on the amount of SMBs present. SMBs and oligonucleotides were incubated for 15 minutes at 25 °C with gentle agitation twice during this period. Reaction tubes were then positioned on the Magnetic Separation Rack for approximately 1 minute to allow collection of SMBs that now had substrate oligonucleotides attached. Supernatant was removed and discarded. SMBs were then resuspended in 25 μ l of 1 x SMB wash buffer containing 500 pmol of free biotin and incubated for 10 minutes with gentle agitation. Reaction tubes were again positioned on the Magnetic Separation Rack for

approximately 1 minute and supernatant was removed. SMBs were then resuspended in 100 μ l of 1 x SMB wash buffer, again positioned on the Magnetic Separation Rack for approximately 1 minute and supernatant was removed. A final wash step using 100 μ l of 1 x reaction buffer was performed again using the Magnetic Separation Rack to separate SMBs and supernatant. These wash steps ensured all free oligonucleotides and free biotin molecules were removed and would not interfere with the subsequent assay steps. An appropriate amount of protein was then added to the SMBs with 3 x SMB reaction buffer and the necessary volume of dH₂O to dilute the reaction buffer accordingly. Varying amounts of BSA were included during protein addition and previous wash steps to reduce background binding. Protein and SMBs were left to incubate in a thermomixer (Eppendorf) at 30 °C, 350 rpm for 15 minutes. Reaction tubes were then positioned back on the Magnetic Separation Rack and the supernatant was removed and stored for later analysis (unbound protein). SMBs were twice washed in 100 μ l 1 x SMB reaction buffer using the Magnetic Separation Rack to separate SMBs from supernatant. To SMBs, 1 x reaction buffer containing 1% SDS was added in a volume equalling the total volume in which protein was initially added. Reaction tubes were heated at 100 °C for 15 minutes to denature bound protein and cause its dissociation from oligonucleotides. Finally, reaction tubes were positioned on the Magnetic Separation Rack and supernatant containing bound protein was removed and analysed.

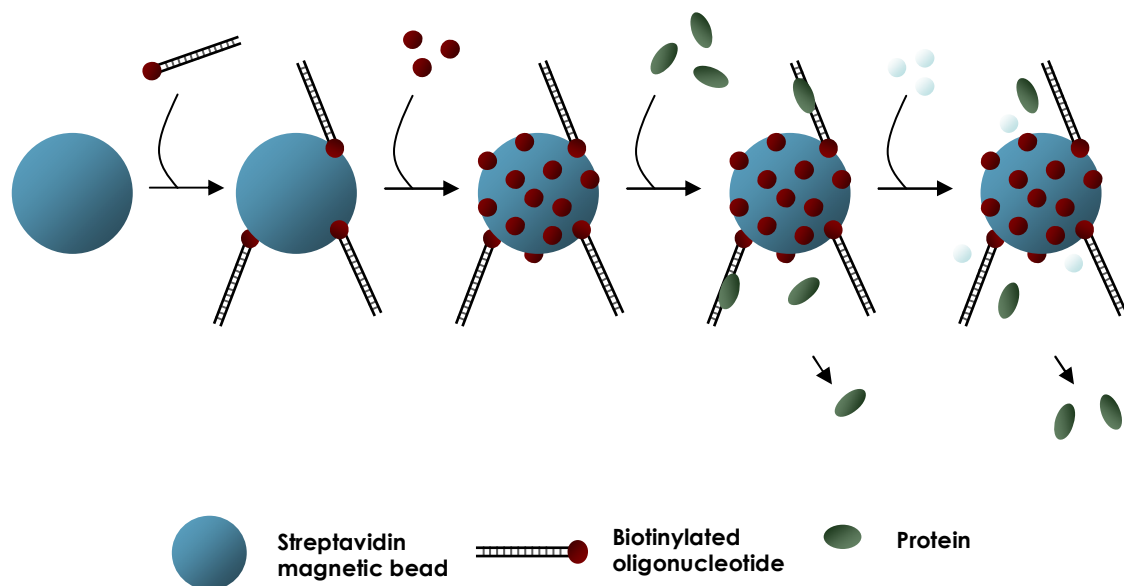


Figure 2.1 *In vitro* SMB DNA binding assay procedure. Biotinylated oligonucleotides are added to SMBs, followed by addition of free biotin molecules (to prevent unspecific protein binding to SMBs). Protein of interest is then added and incubated with oligonucleotides for 15 minutes at 30 °C and 350 rpm. Unbound protein is removed in supernatant following collection of SMB/oligonucleotides/protein with earth magnets. Bound proteins are denatured with SDS, removed in supernatant and analysed with unbound protein removed previously.

Samples collected from DNA binding assays were analysed by SDS-PAGE, western blotting and immunodetection. An image of the membrane was captured using white light on a Universal Hood II (Bio-Rad) and band intensities were measured using the programme 'ImageJ' (National Institutes of Health). By comparing the amount of bound protein and unbound protein from each sample, a percentage of protein binding was calculated. These values can be further used to calculate DNA binding affinities (Section 2.7.3).

2.7.2 SMB Capacity Analysis.

For elucidation of oligonucleotide attachment capacity of SMB (Invitrogen™ Dynabeads® M-280), fluorescently labelled dsDNA and ssDNA (Table 2.7) were added to SMB under the same conditions as an *in vitro* SMB DNA binding assay (Section 2.7.1). Following the 15 minute incubation, beads were twice washed as described and then injected into a 96 well microtitre plate in appropriate volumes of 1 x SMB reaction buffer. To measure fluorescence of samples, KC4 program on an FLx800 microplate fluorescence reader (BioTek®) was used with the following parameters: Excitation – 485 / 20 nm, emission – 516 / 20 nm, optics position - top and sensitivity – 80 nm. Experiments were then performed that analysed fluorescent emission of SMB with either ssDNA or dsDNA attached at various amounts between 0 – 60 pmol.

2.7.3 Dissociation Constant (Kd) Equation used During SMB DNA Binding Assays.

Abbreviations –

DNA = D	Protein = P	Complexed DNA/Protein = DP
Total DNA = Dt	Total Protein = Pt	Concentration of DNA when Protein Binding is 50 % = [Dt, 50]
Free DNA = Df	Free Protein = Pf	

Dissociation constant equations -

Df + Pf ↔ DP		1
[Dt] = [Df] + [DP]	and	[Pt] = [Pf] + [DP]
Kd = $\frac{[Df][Pf]}{[DP]}$	or	$\frac{([Dt] - [DP])([Pt] - [DP])}{[DP]}$

$$\text{At 50\% binding: } [DP] = 0.5 [Pt] \quad \text{or} \quad [Pt] = 2 [DP] \quad 4$$

$$Kd = \frac{([Dt] - [DP]) (2 [DP] - [DP])}{[DP]} = \frac{[DP] ([Dt] - [DP])}{[DP]} = [Dt] - [DP] \quad 5$$

$$Kd = [Dt, 50] - 0.5 [Pt] \quad 6$$

2.8 *In vitro* Protein-G Bead (PGB) Assay.

The *in vitro* PGB assay (Figure 2.2) can detect protein-DNA interactions with various substrate DNA molecules. Unlike the SMB assay also used in this study (which is restricted to using biotinylated DNA), protein binding to any molecule can be investigated. The concept of this assay relies on protein-G coated magnetic beads that can precipitate antibodies specifically bound to a protein of interest that will in turn also pull-down any DNA that it is interacting with. Incubation times and temperatures of this assay vary depending on the protein studied. All described conditions are those used in a standard experiment investigating ScKu proteins.

Initially, an appropriate amount of protein (approximately 400 ng) was incubated with excess specific antibody (typically 450 ng) in a reaction solution that also included 2 µl 500 mM KCl, 2 µl 10 x VP buffer and an appropriate volume of dH₂O (final reaction volumes were 20 µl following DNA addition). This solution was typically incubated at 4 °C for 20 minutes. Following this, an appropriate amount of DNA was added to the solution. The amount of DNA added depended on sensitivity of the technique used to visualise assay results. For oligonucleotides, 60 pmol of DNA was added to each reaction and for plasmid DNA 200 ng was added. The amount of protein used depended on substrate DNA and was always in excess. After DNA addition, samples were thoroughly mixed and incubated on ice for 30 minutes. For certain proteins, DNA and antibody incubations were reversed so protein and DNA interactions could occur prior to addition of the antibody. During protein-DNA incubations, 12.5 µl of Dynabeads[®] Protein-G (Invitrogen[™]) were added to 1.5 ml microcentrifuge tubes and washed 3 times with 100 µl 1 x PGB buffer (using Neodymium rare earth permanent magnet Magnetic Separation Rack (New England BioLabs[®] Inc.) to collect PGBs and remove supernatant each time). Following this, reaction solutions were added to PGBs and incubated at 10 °C at 450 rpm for 30 minutes. The Magnetic Separation Rack was then used to separate PGB/antibody/protein/DNA complexes from the supernatant and this supernatant was removed and kept for analysis (unbound DNA). PGBs were then washed 3 times with 100 µl 1 x PGB buffer. To release bound DNA, 1% SDS was added in a

total volume of 20 μ l and samples were heated at 65 °C (450 rpm) for 5 minutes. Bound DNA was then collected in the supernatant.

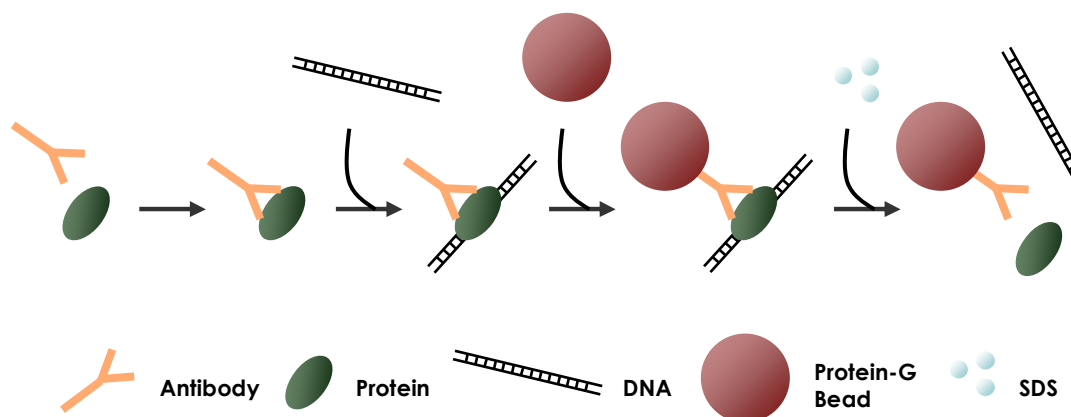


Figure 2.2 *In vitro* PGB DNA binding assay procedure. Antibodies are incubated with specific proteins before substrate DNA molecules are introduced. Protein and DNA is then incubated to allow protein-DNA interactions to occur. PGBs are then introduced and can be used to pull down protein-DNA complexes via attached antibodies. A magnetic force is applied to collect PGB/antibody/protein/DNA complexes and any unbound DNA is removed in the supernatant. Following wash steps, bound DNA is released from protein with 1% SDS and analysed by gel electrophoresis.

For plasmid (as binding substrate) assays, bound and unbound DNA samples were analysed by agarose gel electrophoresis (Section 2.2.6) and band intensities were measured using the programme ‘ImageJ’ (National Institutes of Health) to calculate values for percentage DNA bound. For oligonucleotide assays, bound oligonucleotides were subjected to native polyacrylamide gel electrophoresis (Section 2.3.2), transfer to nylon membranes and chemiluminescent detection of biotinylated tags (Section 2.3.3).

2.9 *In vitro* Co-Immunoprecipitation (Co-IP) Assay.

The production of antibodies specific for ScKu1 and ScKu2 (Section 2.14) provided the opportunity to adapt the protein-DNA binding assay described in Section 2.8 to allow identification of interactions between both Ku proteins. Essentially, the Co-immunoprecipitation (Co-IP) assay is identical to the DNA binding assay except that a secondary protein is added to reactions instead of DNA. Briefly, 600 ng of protein was incubated for 15 minutes at 4 °C with or without 600 ng of a different protein in a 20 μ l reaction volume containing 2 μ l of 1 M KCl, 2 μ l 10 x VP and dH₂O. An appropriate amount of antibody specific for 1 protein was subsequently added and reactions were incubated for a further 10 minutes at 4 °C. Reaction solutions were then added to 12.5 μ l of Dynabeads® Protein-G (Invitrogen™) and incubated at 10 °C for 20 minutes at 450 rpm with occasional

agitation. Non-precipitated protein was collected by removal of the supernatant following positioning of reaction tubes on a Magnetic Separation Rack (New England BioLabs® Inc.) and precipitated protein was collected by denaturing complexes with 15 µl of 1% SDS, heating at 65 °C for 10 minutes, and removing supernatant after application of the magnetic force. Precipitated and non-precipitated protein was analysed by SDS-PAGE and western blotting. Band intensities were measured using 'ImageJ' (National Institutes of Health).

2.10 *In vitro* Ligation Assay.

Linearisation of pUC18 was performed by digestion with either *Eco*R1 (overhang-ends) or *Sma*I (blunt-ends). T4 DNA ligase (Roche) was used to ligate plasmids with or without prokaryotic Ku or BSA present in 20 µl reactions containing: 1 µl T4 DNA ligase, 2 µl 10 x ligation buffer, 200 ng plasmid DNA, an appropriate amount of Ku or BSA proteins and dH₂O. Reactions were incubated at various temperatures for an appropriate length of time with occasional agitation. Ligations were stopped by addition of 6 x loading buffer and stored at 4 °C. Samples were then analysed by agarose gel electrophoresis (Section 2.2.6).

2.11 Analytical Ultracentrifugation.

Analytical ultracentrifugation (AUC) experiments were performed using a Beckman XL1 analytical ultracentrifuge equipped with absorbance optics. Sedimentation equilibration experiments were undertaken at 20 °C and at 9000 or 12000 rpm. Concentration profiles were measured at 260 nm. Scans were recorded every 4 h to determine when equilibrium had been reached. After equilibrium had been achieved, 4 scans were recorded for each sample. AUC experiments and analysis of raw data was carried out by Dr Tom Clarke, UEA.

2.12 Dynamic Light Scattering.

Dynamic light scattering (DLS) experiments were carried out on a DynaPro 99 (Protein Solutions) using 'Dynamics V6' programme for data analysis. Prior to analysis, protein was centrifuged at 5000 rpm at 4 °C for 20 minutes and supernatant used in subsequent analysis. Before protein measurements, 0.2 µM filtered dH₂O was measured to ensure count fluctuations were less than 1000. Protein was in 1 x protein storage buffer and all measurements were performed at 20 °C. For all samples, acquisition times were set to 10 seconds for a total of 20 acquisitions to measure 15 µl of 1 mg / ml protein.

2.13 Chromatin Immunoprecipitation (ChIP).

2.13.1 Cross-linked Chromatin Preparation.

All buffers used in chromatin preparation and ChIP assays were made with nanopure water and contained protease inhibitors.

HeLa cells at 80% confluence (approximately 1×10^7 cells per plate) were provided on 154 cm² plates that had been grown at 37 °C. A number of plates were twice irradiated with 20 J / m² using a UV Stratalinker[®] 1800 (Stratagene[®]) and, following irradiation, plates were grown for a further 15 hours at 37 °C. To fix cells, formaldehyde was added at 1% final concentration to the sides of each plate and then evenly distributed to prevent cell disturbance. Plates were incubated at 25 °C with gentle agitation for 10 minutes. Following this, glycine was added to a final concentration of 0.125 M and was evenly distributed. Media/formaldehyde/glycine was removed and plates were washed with 10 ml ice cold PBS. Cells were scraped into 1 ml PBS containing 1 x protease inhibitors (Complete Mini, EDTA-free, Roche) and centrifuged at 4 °C, 1200 rpm for 5 minutes. The resulting supernatant was removed and the cell pellet was resuspended in 1 x SDS lysis buffer (2-3 x 10⁶ cells / 200 µl 1 x SDS lysis buffer), split into 400-500 µl aliquots and incubated at 4 °C for 10 minutes. Cell lysates were sonicated for 12 x 30 seconds to shear chromatin, with samples kept on ice constantly throughout the procedure. Following sonication, samples were centrifuged at 4 °C, 1300 rpm for 10 minutes. To measure yield of sheared chromatin, absorbance at OD₂₆₀ of 1.5 µl sample was measured using program ND-1000 on a Nano-drop[®] spectrophotometer (Labtech). Chromatin was snap frozen in liquid nitrogen and stored at -80 °C.

2.13.2 ChIP assay.

Approximately 200 µg of prepared cross-linked chromatin fragments were diluted 1:10 in 1 x dilution buffer. Chromatin (50 µg) was set aside for later analysis as 'total chromatin input' sample. Staph A cells (100 µl) were added to samples and incubated at 4 °C with rotation for 30 minutes. Samples were then centrifuged at 2500 rpm for 5 minutes at 4 °C and the resulting supernatant removed and split into two. An irrelevant antibody (5 µg) was added to one half of this sample, and 5 µg of a specific antibody was added to the other half. Samples were incubated overnight at 4 °C with rotation. Staph A cells (50 µl) were then added to both samples and incubated at 4 °C with rotation for 1 hour. Following this, samples were spun at 2500 rpm for 5 minutes at 4 °C. Supernatant was removed and discarded and pellets were resuspended in 1 ml ice cold 1 x low salt buffer followed by rotation for 5 minutes and

centrifugation at 4000 rpm for 1 minute at 4 °C. The supernatant was removed and the process was repeated with 1 x high salt buffer and then 1 x LiCl buffer. Samples were washed twice with 1 x TE buffer and resuspended in 500 µl of 1 x elution buffer (including ‘total chromatin input sample’). Samples were incubated with rotation at 4 °C for 20 minutes and then centrifuged at 10000 rpm for 10 minutes. Supernatant was removed and cross-links were reversed by addition of 20 µl of 5 M NaCl and incubation at 65 °C for 5 hours.

Following cross-link reversal, 10 µl 0.5 M EDTA, 2 µl 1 M Tris-HCl, pH 6.5 and 2 µl of 10 mg / ml proteinase K were added to each sample and incubated at 45 °C for 1 hour. DNA was recovered by addition of 500 µl of phenol to 500 µl of sample and this mixture was vortexed and centrifuged at 13000 rpm for 5 minutes. The top layer was removed and 200 µl of 100% EtOH was added to each sample and incubated at 25 °C for 10 minutes. QIAquick® Gel Extraction Kit (Qiagen) was then used to recover DNA (according to manufacturer’s instructions). DNA was eluted in 40 µl 1 x TE buffer and analysed by PCR (Section 2.2.5).

2.14 Antibody Production for ScKu1 and ScKu2.

1.5 mg of ScKu1 and ScKu2 were used as antigens to produce polyclonal antibodies to be used in future experiments. Antibodies were produced by Cambridge Research Biochemicals®.

2.15 Media and Buffers.

2.15.1 Media.

Table 2.9 Media used for bacterial culture.

Media name	Media constituents
Luria-Bertani broth (LB)	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl.
LB agar	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar.
SOC media	2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 20 mM glucose, pH 7 with NaOH.
Mannitol soya flour medium (SFM)	2% (w/v) agar, 2% (w/v) Mannitol, 2% (w/v) Soya flour.
Tryptone soya broth (TSB)	3% (w/v) Oxoid Tryptone Soya Broth powder (CM129).

2.15.2 Additions to Bacterial Culture Media.

All additional constituents of media were sterilised using a 0.2 µm filter. The antibiotics used were: Ampicillin (final concentration 50 µg/ml), kanamycin (final concentration 50 µg/ml), chloramphenicol (final concentration 34 µg/ml), apramycin (final concentration 50 µg/ml) and hygromycin (final concentration 50 µg/ml). For induced expression of recombinant proteins, IPTG was added at a final concentration of 0.4 mM.

2.15.3 Buffers Used in DNA Procedures.

Table 2.10 Buffers used during analysis of DNA.

Buffer name	Buffer constituents
1x Tris-acetate-EDTA (TAE)	40 mM Tris-acetate pH 7.8, 1 mM EDTA.
1 x Tris-EDTA (TE)	10 mM Tris-HCl pH 8, 1 mM EDTA.
1 x Tris-borate-EDTA (TBE)	90 mM Tris-borate, 2 mM EDTA.
6 x native loading buffer	10 mM Tris-HCl pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol, 60 mM EDTA.
1 x stop solution	20 ml deionised formamide, 400 µl 0.5 M EDTA, 40 µl 1 M NaOH, 20 mg bromophenol blue.

2.15.4 Buffers Used during Protein Procedures.

Table 2.11 Buffers used in protein purification and analysis.

Buffer name	Buffer constituents
8 x binding buffer for His·Bind® column / HisTrap™ HP cartridges	4 M NaCl, 160 mM Tris-HCl pH 7.9, 40 mM Imidazole.
8 x charge buffer for His·Bind® column / HisTrap™ HP cartridges	400 mM NiSO ₄ .
8 x wash buffer for His·Bind® column / HisTrap™ HP cartridges	4 M NaCl, 160 mM Tris-HCl pH 7.9, 480 mM Imidazole.
8 x elution buffer for His·Bind® column / HisTrap™ HP cartridges	2 M NaCl, 80 mM Tris-HCl pH 7.9, 4 M Imidazole.
p53 purification buffer A	10 mM HEPES, pH 7.4, 1.5 mM MgCl ₂ , 5 mM KCl, 1 mM DTT, protease inhibitors.
p53 purification buffer B	10 mM HEPES, pH 9, 1.5 mM MgCl ₂ , 5 mM KCl, 10 mM DTT, protease inhibitors.
p53 purification buffer B2	10 mM HEPES, pH 9, 1.5 mM MgCl ₂ , 200 mM KCl, 10 mM DTT, protease inhibitors.
1 x Tris-glycine SDS (TGS)	25 mM Tris pH 8.3, 192 mM glycine, 0.1% (w/v) SDS.
4 x SDS-PAGE loading buffer	200 mM Tris-HCl pH 6.8, 8% (w/v) SDS, 0.4% (w/v) bromophenol blue, 40% (w/v) glycerol, 400 mM β-mercaptoethanol.

Coomassie blue	0.25 g Coomassie Brilliant Blue R-250, 90 ml methanol: H ₂ O (1:1, v/v), 10 ml glacial acetic acid.
1 x electrophoresis blotting buffer	24 mM Tris-HCl pH 7.5, 192 mM glycine, 20% methanol.
1 x Tris-buffered saline (TBS)	50 mM Tris-HCl pH 7.5, 150 mM NaCl.
1 x Tris buffered saline + Tween (TBST)	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween [®] 20.
1 x Phosphate buffered saline (PBS)	10 mM Phosphate-HCl pH 7.4, 137 mM NaCl, 2.7 mM KCl.
1 x Phosphate buffered saline + Tween (PBST)	10 mM Phosphate-HCl pH 7.4, 137 mM NaCl, 2.7 mM KCl, 0.1% Tween [®] 20.
1 x protein storage buffer	20 mM Tris-HCl pH 7.5, 100 mM NaCl.

2.15.5 Buffers used during *in vitro* Assays.

Table 2.12 Buffers used during *in vitro* DNA binding and ligation assays.

Buffer name	Buffer constituents
1 x SMB wash buffer	500 mM NaCl, 20 mM Tris-HCl pH 7.5, 1 mM EDTA.
1 x SMB reaction buffer	10 mM Tris-HCl pH 7.5, 50 - 250 mM KCl, 1 mM DTT, 200 - 400 µg/ml BSA.
3 x SMB reaction buffer	30 mM Tris-HCl pH 7.5, 150 - 750 mM KCl, 3 mM DTT, 600 - 1200 µg/ml BSA.
10 x VP buffer	50 mM Tris-HCl pH 7.8, 0.1% Triton X-100.
1 x PGB buffer	5 mM Tris-HCl pH 7.8, 50 - 250 mM KCl, 0.1% Triton X-100.
10 x ligation buffer	660 mM Tris-HCl pH 7.5, 50 mM MgCl ₂ , 10 mM DTT, 10 mM ATP.

2.15.6 Buffers Used during *in vivo* Assays.

Table 2.13 Buffers used during *in vivo* experiments.

Buffer name	Buffer constituents
1 x SDS lysis buffer	50 mM Tris-HCl pH 8.1, 1% SDS, 10 mM EDTA.
1 x dilution buffer	16.7 mM Tris-HCl pH 8.1, 167 mM NaCl, 0.01% SDS, 1.2 mM EDTA, 1.1% Triton X-100.
1 x low salt buffer	20 mM Tris-HCl pH 8.1, 150 mM NaCl, 0.1% SDS, 2 mM EDTA, 1% Triton X-100.
1 x high salt buffer	20 mM Tris-HCl pH 8.1, 500 mM NaCl, 0.1% SDS, 2 mM EDTA, 1% Triton X-100.
1 x LiCl wash buffer	10 mM Tris-HCl pH 8.1, 0.25 M LiCl, 1 mM EDTA, 1% NP40, 1% deoxycholate.

CHAPTER 3

Characterising Human p53-DNA Interactions

As a central component of several complex networks within multicellular organisms, p53 is undoubtedly integral to the maintenance of genomic stability. Renowned for its activities as a transactivator of genes involved in cell cycle regulation, DNA repair and apoptosis, it has also recently been proposed that p53 may function as part of a global surveillance mechanism responsible for detecting and responding to genomic regions prone to potentially detrimental events (Walter et al., 2005). Interactions with DNA form an integral element of the function of p53 and it is recognised that it can associate with DNA in diverse and versatile modes depending on both the DNA substrate and the status of p53 itself (Kim and Deppert, 2006). Interactions between p53 and consensus response elements (REs) in target gene promoters have been well characterised (Shu et al., 2006). However, p53 has also been shown to bind to DNA in a manner strongly dependent on DNA structure and also non-specifically to linear dsDNA (McKinney et al., 2004).

Recently, secondary structures that form in TRS tracts have been suggested as natural p53 binding sites (Walter et al., 2005). Such activity might have considerable physiological implications and ultimately may influence characteristic phenotypes of some human hereditary neurological disorders. Although it has been documented that latent p53 demonstrates a strong binding affinity towards hairpin structures composed of CTG•CAG repeats, the consequences of such interactions remain unknown. It has been proposed that as this activity is exhibited by unmodified p53, this may be evidence reinforcing the hypothesis that unmodified p53 is active as a regulator of genome stability.

The *in vitro* DNA binding activity of p53 has been intensely investigated, often by using electrophoretic mobility shift assays (EMSAs) to analyse binding patterns of p53 to various DNA substrates. As a consequence of these studies, models were proposed inferring that p53 required modifications or even complete removal of the CTD for initiation of sequence-specific binding (Anderson et al., 1997; Hupp and Lane, 1994), a concept that was later disproven as the CTD was actually found to enhance sequence-specific binding by associating p53 close to DNA. As a considerable amount of the current knowledge of the DNA binding activity of p53 has derived from EMSAs, it would be advantageous to develop alternative

DNA binding assays for further characterisation of p53-DNA interactions and for verification of previously obtained data. As doubt exists regarding the sensitivity of EMSAs during the detection of weak protein-DNA interactions (as may be the case for non-sequence specific binding), assays that are sensitive enough to detect and differentiate between different modes of these interactions would be highly beneficial.

A main aim of this section of work was to optimise a recently developed *in vitro* streptavidin magnetic bead (SMB) DNA binding assay for use as a technique for investigating human p53 interactions with DNA. Initially, 2 types of human p53 were analysed – a full length version that had been over-expressed and extracted from eukaryotic insect cells, and a full length recombinant version that had been over-expressed and purified via a polyhistidine tag from prokaryotic *E. coli*. As modifications have previously been implicated to influence p53-DNA binding, it was of considerable interest to ascertain if over expression in prokaryotic or eukaryotic cells affected DNA binding properties. Prior to their use in SMB assays, oligonucleotides that would ultimately act as DNA binding substrates required production and subsequent analysis to ensure the desired products had been successfully formed. It was also crucial to analyse the SMBs employed during assays to quantify elements that could potentially interfere with assay results, including oligonucleotide attachment capacity and unspecific protein binding. Simultaneously, Protein-G beads (PGB) that form components of a separate *in vitro* DNA binding assay were also examined alongside SMBs to look for similar properties. After analysis of SMB and PGB components, p53 DNA binding assays were performed to investigate precise activities of both types of p53, initially to determine optimum conditions for subsequent assays such as the effect ionic strength has on p53-DNA interactions. Following comparisons of both p53s that were available, assays focused on interactions with unusual DNA structures, predominantly to hairpin regions, and identification of whether interactions were dependent upon structure or sequence.

Lastly, *in vivo* DNA binding properties of human p53 were also investigated by utilising chromatin immunoprecipitation (ChIP) assays that have been developed for use in HeLa cells. Binding to known p53 REs in Mdm2 and p21 promoter regions were analysed alongside binding to GAPDH promoter (a region believed to not be a p53 target binding site). DNA binding by p53 was investigated in cells that had or had not been subjected to UV irradiation to identify if DNA damage influenced p53-DNA binding.

3.1 Human p53 Protein Analysis.

During this study, DNA binding activities of 2 versions of human p53 were investigated. To assess the influences that post-translational modifications can exert on p53 activity, it was of interest to identify if human p53 derived from bacterial host cells demonstrated comparative activity to human p53 over-expressed in insect cell models. Insect cells are generally accepted to contain advanced cellular machinery that can facilitate proper folding of eukaryotic-derived proteins and allow complex post-translational modifications of mammalian proteins. By contrast, such modifications may not be possible in prokaryotic systems, and this can, ultimately, influence protein activity. High Five insect cells (Invitrogen™) that had been infected with recombinant baculovirus directing expression of human wild-type p53 (p53-I) were kindly supplied by Ella Kim (University of Lübeck, Germany). As p53 expression was directly controlled by baculoviral transcription, cellular levels of recombinant p53 were significantly greater than other cellular proteins allowing quick extraction as described (Section 2.4.3.6), a method widely used during p53 studies (Dudenhoffer et al., 1998; Gohler et al., 2002; Walter et al., 2005). A recombinant human wild-type p53 protein (Active Motif®) that had been over-expressed in *E. coli* was also investigated (p53-B). This recombinant protein had an additional N-terminal polyhistidine tag that allowed purification via an affinity column in combination with fast performance liquid chromatography prior to this study.

3.1.1 Insect Cell Derived Human p53 Extraction and Analysis.

Prior to p53-DNA interaction assays, it was essential to ensure samples contained the expected p53 protein. Human p53 was extracted from insect cells and analysed via SDS-PAGE and Coomassie staining (Figure 3.1 A). Analysis revealed p53 had successfully been extracted as a relatively intense band present at approximately 50 kDa, however a considerable amount of unidentified cellular proteins were also present. A Bradford assay was performed, which estimated the total protein concentration to be approximately 1 µM, but, as this encompassed all proteins extracted, the actual concentration of p53 was significantly less than this.

To confirm the presence of human p53 in the extracted p53-I sample, western blots were performed followed by immunodetection analysis using p53 (DO-1) (Santa Cruz) antibody, specific for amino acids 11-25 of human p53. Western blots identified that the optimum p53 (DO-1) dilution for analysing 5 – 0.1 pmol of p53-I was 1:100,000, indicating a high

sensitivity of the antibody for human p53. Two samples of p53-I that had been extracted separately were analysed at varying amounts (Figure 3.1 B and C) to discover an appropriate amount that would produce bands of relatively high intensity without antibody binding non-specifically to contaminating proteins.

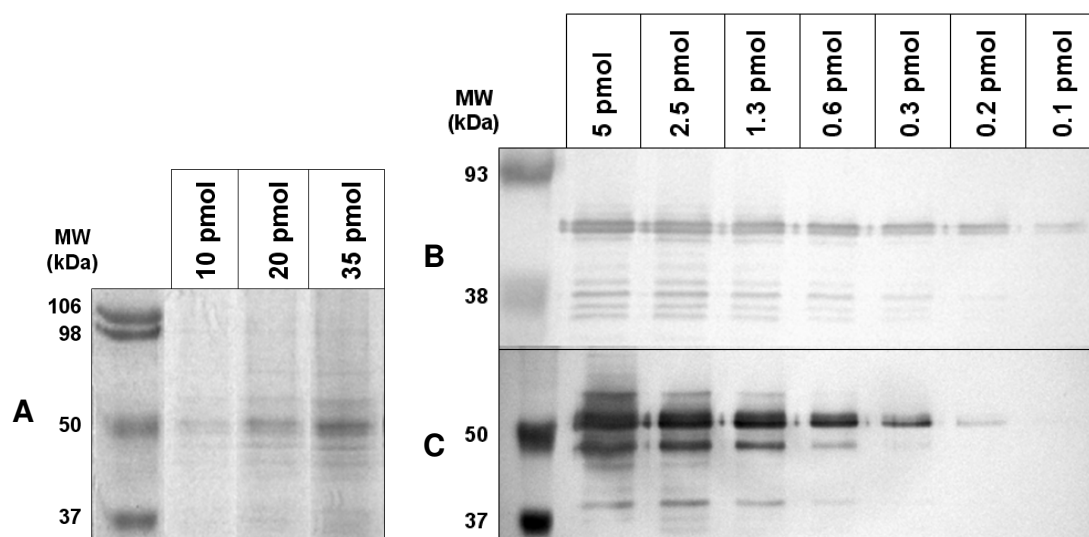


Figure 3.1 Electrophoretic analysis of human p53 derived from insect cells. **A.** SDS-PAGE (using 10% polyacrylamide gels) followed by Coomassie staining of p53-I. The relatively intense band at approximately 50 kDa indicates p53 was successfully obtained, however significant levels of contaminating proteins were present. **B.** SDS-PAGE followed by western blot analysis using p53 (DO-1) at 1:100,000, confirming the intense band detected in (A) was human p53. **C.** SDS-PAGE followed by western blot analysis using p53 (DO-1) at 1:100,000 of a separate extraction of p53-I.

Decreasing amounts of p53-I to 1.5 – 0.2 pmol resulted in clear bands representing p53, with considerably less unknown bands either caused by non-specific immunodetection or the presence of p53 breakdown products. It was noticed that the p53 band appeared to consist of 2 bands of similar size that were detected to a similar extent by p53 (DO-1), except for p53-I shown in Figure 3.1 C, which appears to have a more intense lower band. The reason behind detection of two bands is unknown, however it may be possible these samples contain multiple species of p53 exhibiting alternative modifications that may influence electrophoretic mobility. It was accepted that for p53-I, amounts varying between 1.5 – 0.2 pmol would be acceptable for future western blot analysis.

3.1.2 Human Recombinant p53-B Analysis.

As mentioned, an alternative recombinant human p53 protein that had been over-expressed in *E. coli* (p53-B) was also analysed during this study. Coomassie staining had already been

performed by the manufacturer, showing the p53 samples were relatively pure (Appendix I). The addition of an N-terminal polyhistidine tag was utilised for western blot analysis using anti-His antibody, in addition to immunodetection of the p53 protein using an anti-p53 antibody, which both show the observed band was human p53. Immunodetection performed by the manufacturer with the anti-p53 antibody suggests the presence of contaminating proteins or p53 breakdown products, however this is probably a consequence of using the anti-p53 at an inappropriate dilution. It was of interest to analyse p53-B using p53 (DO-1) as this was the antibody selected for p53-I analysis and would be used in future western blots. As before p53 (DO-1) was used at 1:100,000 to analyse 0.5 pmol of both p53-B and p53-I (Figure 3.2).

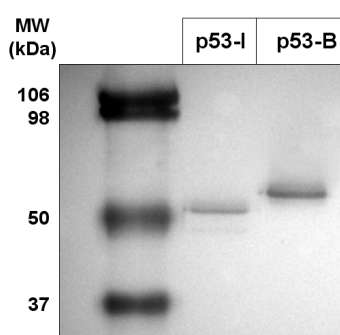


Figure 3.2 Electrophoretic and western blot analysis of p53-I and p53-B using p53 (DO-1) during immunodetection. Both p53-I and p53-B (0.5 pmol) were subjected to SDS-PAGE using 10% polyacrylamide gels and western blotting followed by immunodetection with p53 (DO-1) antibody at 1:100,000.

The results obtained from this analysis were unexpected. Both p53 samples contained a band at approximately 50 kDa, however p53-B demonstrated reduced migration compared to p53-I, indicating it had a greater molecular weight. As both p53-B and p53-I were full-length versions of p53 it was expected electrophoretic mobility would be similar. Differences in migration may be explained - at least partially - because of the additional 14 amino acids that comprise the polyhistidine tag required for purification. The molecular weight of the histidine residues alone may not fully justify the difference observed, but residual structural features in this region that may not have been fully denatured may also retard migration. An alternative explanation is that p53-B and p53-I may consist of alternate splice variants of p53. It is known that alternative transcripts of p53 exist in humans and up to 10 p53 isoforms have been identified that can act synergistically or antagonistically depending on their conformation (Marcel and Hainaut, 2009). If p53-B and p53-I are indeed different p53 isoforms, it is feasible that both samples would demonstrate varied DNA binding characteristics during *in vitro* DNA binding analysis.

3.2 Oligonucleotide Annealing and Analysis.

For protein-DNA interaction investigations with the *in vitro* SMB assay, various oligonucleotide substrates were produced that would ultimately act as potential protein-binding substrates. The majority of oligonucleotides used for DNA binding assays would be double-stranded, meaning relevant single-stranded oligonucleotides had to be annealed together to produce the desired products. All substrates required at least 1 biotinylated strand that would attach to Dynabeads® M-280 Streptavidin-coated beads during the assays. Following the annealing procedure, conformations of the resulting substrates were analysed to ensure the appropriate molecules had been produced.

DNA conformation analysis can be performed using a variety of procedures, including NMR and dynamic programming algorithms (SantaLucia and Hicks, 2004). During this study, oligonucleotide conformations were investigated using electrophoretic mobility analysis by native and denaturing polyacrylamide gel electrophoresis, another widely used technique for DNA structure analysis (Chastain et al., 1995; Cooper and Hagerman, 1987; Duckett et al., 1988; Kahn et al., 1994). This method relies on the reduced migration of larger DNA molecules or DNAs containing unusual structural features that restricts migration during electrophoresis. As various DNA substrates of different sizes and conformations were employed during *in vitro* SMB assays, migration of these molecules could be directly compared to each other to allow identification of successful annealing.

A proportion of DNA substrates used in these studies contained internal secondary structures and it was crucial to confirm the structures had been successfully incorporated to prevent unreliable data from the subsequent assays. Hairpin structures were introduced into selected oligonucleotides by annealing together 2 ssDNAs that had strong base pairing at the substrate termini and with 1 strand containing excess bases in the central region. Introducing central hairpins to oligonucleotides significantly alters the topology of the DNA molecules, and a resulting extended 'Y-shaped' conformation with pseudo-threefold symmetry is usually obtained (Altona et al., 1996). The stability of the hairpin is dependent upon the base pairing either side of the hairpin. Depending on factors such as protein interactions and ionic strength, topology of the DNA can alter and folding and stacking of helices may occur, which can significantly affect electrophoretic mobility.

Alongside various sized dsDNA, ssDNA and dsDNA containing internal hairpin structures, double-stranded oligonucleotides containing alternative features including mismatches and

base pair insertions were also produced as these characteristics can also significantly affect protein-DNA interactions.

3.2.1 Optimisation of Annealing Procedure.

Preliminary annealing procedures involved the addition of relevant single-stranded oligonucleotides together in equal amounts, heating at 100 °C for 8 minutes and then allowing the DNAs to slowly cool to 20 °C for approximately 6 hours. The resulting dsDNA oligonucleotides were analysed by native polyacrylamide gel electrophoresis (Section 2.3.2) and visualised by chemiluminescent detection (Section 2.3.3). Only biotinylated DNA was recognised during this detection, but, as only biotinylated DNA affected the *in vitro* SMB DNA binding assays, this was adequate.

Initial experiments focused on the production of double-stranded B-DNAs [p53 p21-RE], [p53 NBS1] and a dsDNA containing an internal hairpin - [p53 (CTG)₇ HP] (all oligonucleotide substrates shown in Table 2.7). Analysis of annealing revealed a high proportion of biotinylated ssDNA molecules present in the dsDNA samples as 2 bands representing dsDNA and ssDNA were present. This was confirmed following denaturing electrophoresis of the same samples when only single bands, representing ssDNA, were observed that had migrated identically to the lower bands present in native electrophoresis analysis. However, since biotinylated ssDNA contamination of dsDNA samples could have a considerable effect on *in vitro* SMB DNA binding experiments, the annealing procedure for producing DNA substrates required optimisation.

Increasing environmental ionic strength can facilitate stabilisation of nucleic acid structures. The highly negative phosphate backbone of DNA can destabilise the structure of the molecule (Oda and Nakamura, 2000) and can negatively effect dsDNA formation. Addition of salt effectively neutralises the negative repulsive charges and enhances stabilisation of the dsDNA conformation. In order to facilitate annealing of dsDNAs, various concentrations of KCl between 50 – 250 mM were added prior to annealing. It was determined that increasing KCl to 50 mM greatly improved annealing but that subsequent increases in concentration above this value had negligible effects to the observed improvement. Thus, annealing efficiency was improved and contaminating biotinylated ssDNA was reduced (data not shown), however considerable amounts still remained.

To further reduce contaminating biotinylated ssDNA, the use of alternative ratios of biotinylated ssDNA to unbiotinylated ssDNA oligonucleotides were investigated during

annealing. Ratios of 1:2 and 1:4 were selected (Figure 3.3) and it was expected a higher amount of biotinylated ssDNA would become incorporated into dsDNA molecules due to the large excess of unbiotinylated ssDNA. Although the resulting samples would contain large amounts of unbiotinylated ssDNA, this would not interfere with the *in vitro* SMB DNA binding assay as they would be removed during washing procedures. Amounts of biotinylated ssDNA were significantly decreased in dsDNA oligonucleotide samples at a ratio of 1:2 of biotinylated:unbiotinylated and almost all contaminating biotinylated ssDNA were removed using a 1:4 ratio. Thus, optimum annealing conditions had been identified and all double-stranded oligonucleotide samples investigated during the *in vitro* SMB DNA binding assays were produced using a 1:4 ratio of biotinylated:unbiotinylated DNA and at an ionic strength of 50 mM KCl.

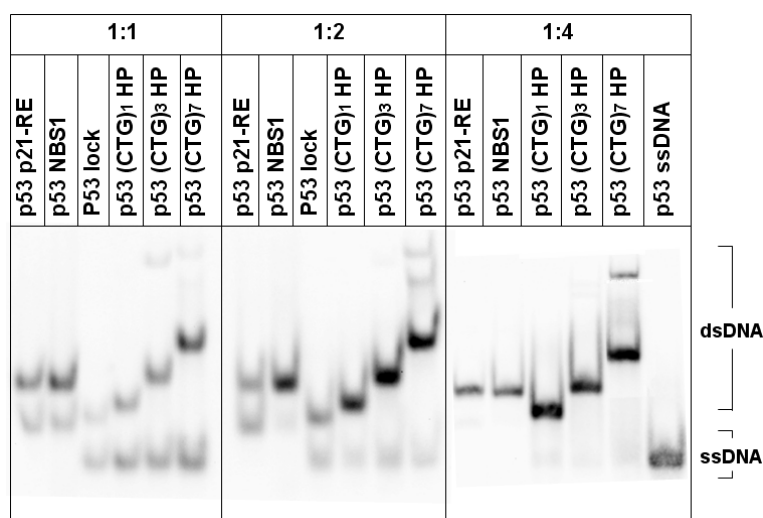


Figure 3.3 Oligonucleotide analysis following annealing of various dsDNAs under different conditions. Initial annealing analysis identified significant amounts of residual biotinylated ssDNA in dsDNA samples. Addition of 50 mM KCl improved annealing efficiency when at a ratio of 1:1 (biotinylated:unbiotinylated). Annealing was improved further by increasing this ratio to 1:2 and annealing was most efficient at a 1:4 ratio. Successful incorporation of hairpins in specific dsDNAs was ascertained by a significant decrease in migration. Sequences and structures of these oligonucleotides are described in Table 2.7 and dsDNA and ssDNA is indicated.

As well as illustrating the improvements in annealing, the results also confirmed that DNAs were in the desired conformations as relative dsDNA migration was as anticipated. Migration through the gel for [p53 p21-RE] and [p53 NBS1] was considerably less than for [p53 ssDNA], which was expected as these substrates are double-stranded and they also contain more nucleotides in each strand. The [p53 (CTG)₁ HP] molecule migrated further than both [p53 p21-RE] and [p53 NBS1]. This oligonucleotide did contain a 3 base bulge that may reduce mobility, but it also contained 21 less nucleotides that will significantly increase migration. Migration of [p53 (CTG)₃ HP] was similar to [p53 p21-RE] and [p53 NBS1], and

considerably less than [p53 (CTG)₁ HP], even though the net increase in nucleotides was only 6. This implied the hairpin secondary structure had successfully been incorporated and was retarding mobility. This assumption is reinforced by [p53 (CTG)₇ HP] migration as mobility was notably less than [p53 p21-RE1] and [p53 NBS1], even though this molecule contained 3 less nucleotides than the dsDNAs without a hairpin. As predicted, the introduction of internal hairpin structures significantly retarded electrophoretic mobility. Interestingly, [p53 (CTG)₇ HP] contained an additional unusual band that showed dramatically retarded migration. Similar bands have been identified in previous studies using this molecule and it was believed to be conformational isomer of the dsDNA containing an internal hairpin (Walter et al., 2005). In this previous study, this idea was reinforced as the slower migrating band also appeared following re-analysis of DNA that had been initially isolated from a single band representing the expected DNA molecule.

3.2.2 Production of Unusual Double-stranded Oligonucleotides.

For protein-DNA binding experiments, various DNA substrates containing different structural features were produced using the optimised annealing procedure described in Section 3.3.1. Following annealing, all samples were analysed by native gel electrophoresis and results are shown in Figure 3.4.

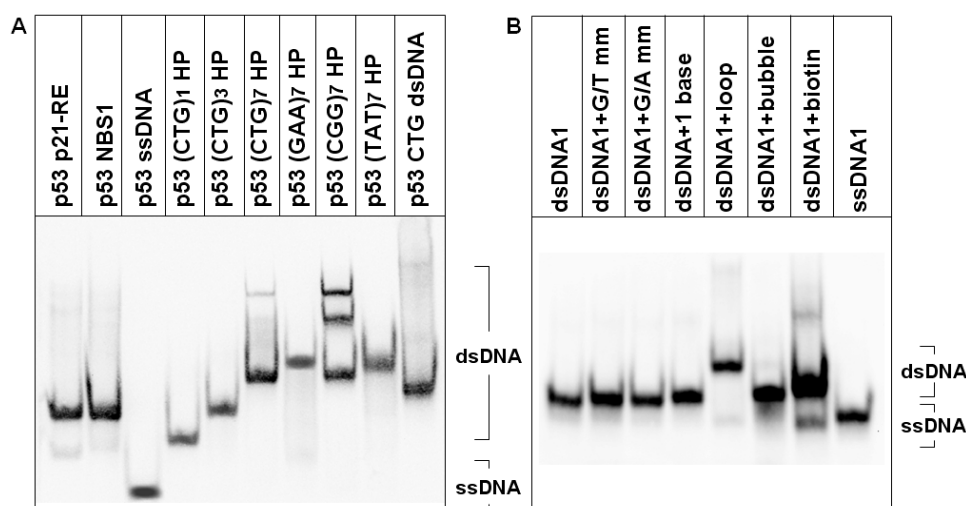


Figure 3.4 Analysis of oligonucleotides annealed using optimised conditions. Following annealing, oligonucleotides were analysed by native gel electrophoresis and biotinylated molecules were detected by chemiluminescent detection. Confirmation of successful annealing was achieved by analysis of the relative migration of each molecule. Sequences and structures of these oligonucleotides are described in Table 2.5 and dsDNA and ssDNA is indicated. **A.** Oligonucleotides used during p53 binding analysis. **B.** Oligonucleotides used during p53 and Ku DNA binding assays.

As shown in Figure 3.4 A, all dsDNA containing internal hairpins consisting of 7 triplet repeats had reduced mobility compared to larger double-stranded B-DNA molecules. Slight variations in migration of these hairpin substrates resulted from alternate sequences in the hairpin regions. Oligonucleotides [p53 (CTG)₇ HP] and [p53 (CGG)₇ HP] migrated further than [p53 (GAA)₇ HP] and [p53 (TAT)₇ HP], potentially resulting from the higher amount of guanine and cytosine nucleotides providing stronger base-pairing, thereby forming a more compact hairpin structure that would encounter less resistance during electrophoresis. Similar to [p53 (CTG)₇ HP], [p53 (CGG)₇ HP] contained a separate band representing DNA that has a very high level of retarded mobility in addition to the band representing the expected structure. These additional DNA molecules could potentially be due to an alternative folding arrangement (Altona et al., 1996). It is unclear why [p53 (GAA)₇ HP] and [p53 (TAT)₇ HP] do not contain structural alternatives but this may be a direct influence of the stronger guanine and cytosine bonds that can form within the hairpin stem of the other DNAs that are more efficient at trapping altered structures. Confirmation of retarded migration by internal hairpins was achieved by direct comparison with [p53 CTG dsDNA]. This dsDNA consists of 96 nucleotides, but it still migrated further than the hairpins consisting of 7 triplet repeats that have 21 fewer nucleotides. These results strongly suggested the key hairpin structures had been introduced successfully.

Figure 3.4 B represents oligonucleotides containing structures that result in less perceivable alterations in electrophoretic mobility than those previously described, as features such as mismatches and single base insertions are more difficult to distinguish than larger hairpin structures. By comparison to [dsDNA1], [dsDNA1+ loop] had reduced migration, implying the presence of an unusual structure; [dsDNA+ bubble] also had slightly reduced mobility when compared to [dsDNA1], which is likely to result from a series of mismatches that perturb the overall structure. Unlike other dsDNA substrates, [dsDNA1+ biotin] contained large amounts of biotinylated ssDNA. This was unavoidable as both ssDNAs prior to annealing were biotinylated and so both were added in 1:1 ratio and not the optimised 1:4 ratio. It also became evident following this analysis that biotin has a significant effect on mobility. This can be seen by comparing [dsDNA1] and [dsDNA1+ biotin]. The extra biotin molecule retarded mobility to a slightly greater extent to that of [dsDNA+ bubble] even though it contains the same number of bases as [dsDNA1]. This idea is also supported by the relative lack of separation between dsDNAs and ssDNAs also evident. The lack of separation between dsDNA and ssDNA can be explained as the addition of biotin is proportionally greater upon smaller ssDNA than larger dsDNA.

3.3 Dynabeads® M-280 Streptavidin and Dynabeads® Protein-G Analysis.

Two alternative approaches for investigating *in vitro* protein-DNA interactions were employed during this study that are dependent on utilising specialised magnetic beads to separate protein-DNA complexes from non-interacting molecules. For the SMB assay (Section 2.7), Dynabeads® M-280 Streptavidin superparamagnetic beads (2.8 µm diameter) were used to capture biotinylated oligonucleotides that would act as substrates for added proteins. These beads have a streptavidin monolayer covalently coupled to the surface to which biotin will attach with high-affinity (Holmberg et al., 2005). This interaction has a dissociation constant of approximately 4×10^{-14} M, making it the strongest non-covalent biological interaction known. For the PGB assay (Section 2.8), Dynabeads® protein-G superparamagnetic beads were used. These beads again were 2.8 µm in diameter, but have a recombinant group G protein from *Streptococcus* covalently coupled to the surface instead of streptavidin. Protein-G can capture protein-specific immunoglobulins that allow attachment of the protein of interest that would ultimately act as substrate for added DNA.

3.3.1 Dynabeads® M-280 Streptavidin Analysis.

3.3.1.1 Capacity Analysis.

For precise measurements of protein-DNA binding affinity towards various substrate DNAs, it is fundamental to identify the amount of DNA that will become physically attached to a set amount of superparamagnetic beads. For M-280 Streptavidin coated beads, the manufacturer states approximately 200 pmol of single-stranded oligonucleotides will attach to 1 mg of Dynabeads, and that capacity is inversely related to DNA molecule size. No values are supplied for double-stranded oligonucleotides, however steric hindrance may reduce binding capacity for larger DNA fragments.

To assess oligonucleotide capacity of SMBs, fluorescently labelled ssDNA (ssDNA+ fluorophore) and dsDNA (dsDNA+ fluorophore) (Table 2.7) were incubated with a set amount of beads under conditions used in the SMB assay and any unattached oligonucleotides were removed in subsequent wash steps. The amount of oligonucleotides attached to beads was determined by fluorescence emission. Initially, it was determined if the presence of SMBs affected fluorescence emission of known amounts of fluorescently labelled DNA. For the amount of beads analysed (10 µl), fluorescence emission from 0 – 30 pmol labelled DNA was not significantly affected with or without beads (data not shown). Consequently, oligonucleotides did not have to be removed from beads during capacity analysis

experiments. To determine variation of fluorescence between labelled ssDNA and dsDNA, fluorescence from varying amounts of both DNAs was analysed (Figure 3.5) in wells that also contained identical amounts of SMBs. The readings suggested that at low amounts there is little difference in fluorescence between DNAs, but at higher amounts dsDNA emitted higher levels of fluorescence.

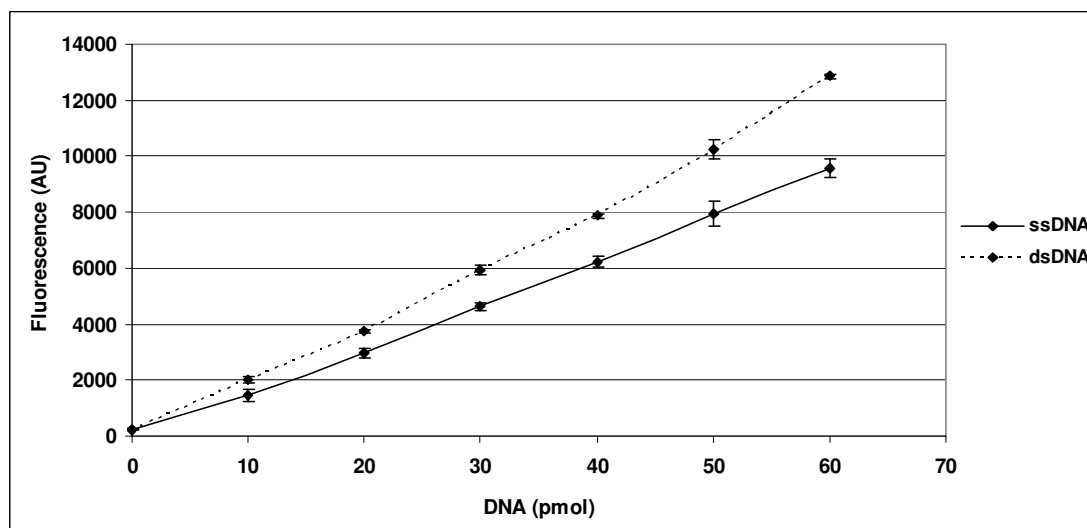


Figure 3.5 Fluorescence emission from ssDNA and dsDNA oligonucleotides. Fluorescence from 0 – 60 pmol of both DNA species was measured using a FLx800 microplate fluorescence reader. Samples contained the relevant DNA, 10 μ l of M-280 Streptavidin-coated beads and an appropriate amount of dH₂O for a final volume of 180 μ l. Averages were taken from at least 3 readings. Standard errors are shown.

The difference in emission between DNA species meant direct comparisons between ssDNA and dsDNA during capacity analysis would not be possible. The reasons for the observed difference was unknown, however variations in DNA concentrations would directly influence fluorescence. For accurate quantification of bead capacity for DNA, calibrations of known amounts of DNA (including beads) were measured simultaneously to experiments analysing bead capacity. It was predicted that as approximately 200 pmol of single-stranded oligonucleotides attaches to 1 mg of SMBs (10 mg / ml), that 20 pmol ssDNA would attach to 10 μ l. For capacity analysis, 10 μ l of beads were prepared and varying amounts of both ssDNA and dsDNA (0 – 60 pmol) were incubated as described (Section 2.7). Importantly, bead-DNA complexes were washed twice to ensure unattached excess DNA was removed. Fluorescence emission from these samples was then measured alongside DNA removed during wash steps and calibration samples of known DNA amounts. Values from calibration samples were used to calculate the amount of DNA attached to beads and excess DNA that did not attach (Figure 3.6 A and B).

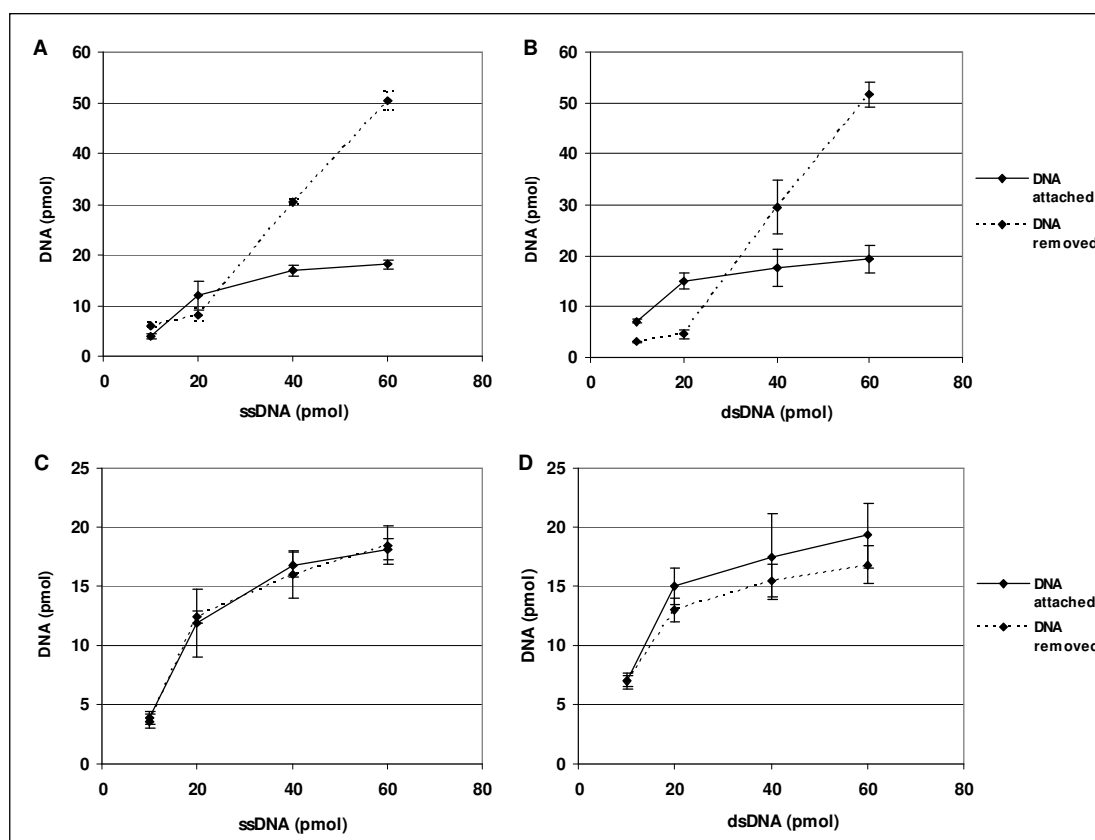


Figure 3.6 Fluorescence emission from labelled DNA species during M-280 Streptavidin-coated bead capacity analysis. [ssDNA+ fluorophore] and [dsDNA+ fluorophore] (Table 2.7) were used during analysis. Averages were taken from at least 4 experiments. Standard errors are shown. **A** – ssDNA attached to beads and excess DNA removed during wash steps. **B** – dsDNA attached to beads and excess DNA removed during wash steps. **C** - Comparing remaining ssDNA attached to beads following wash steps with the amount attached at the primary reading. **D** - Comparing remaining dsDNA attached to beads following wash steps with the amount attached at the primary reading.

It was identified that although the capacity of SMBs was similar to the manufacturer's value, to ensure the expected amount of oligonucleotide attached, the DNA had to be added in excess, preferably by 3 times of the expected amount. This is demonstrated as when 20 pmol of either DNA species is added to 10 μ l beads, approximately only 12-15 pmol becomes attached; however, when 60 pmol of DNA is added almost 20 pmol becomes attached. The results also indicate there is no significant difference between attachment of ssDNA and dsDNA, indicating the level of hindrance caused by the larger dsDNA molecule is low ($t = 0.97$ ($p > 0.05$)) (60 pmol values). It was of interest to discover if wash steps following bead-DNA incubation disturbed DNA that was attached to SMBs. Due to the high-affinity between biotin and streptavidin it was predicted that any loss of DNA would be low. Figure 3.6 C and D shows that although ssDNA fluorescence emission remained almost identical to the pre-wash reading, dsDNA demonstrated a degree of reduced emission in samples with higher amounts of DNA.

3.3.1.2 Non-specific Protein-bead Interactions.

Non-specific protein binding to assay components is an important factor to address during immunoprecipitation experiments. To analyse non-specific protein binding to M-280 Streptavidin-coated beads, *in vitro* SMB assays with and without DNA were performed to detect non-specific p53 binding to beads. Assays were performed according to Section 2.7, using 12 µl of SMBs that were expected to allow approximately 24 pmol of substrate DNA to attach and act as substrate for 1.5 pmol of p53. DNAs analysed included [p53 p21-RE], [p53 NBS1] and [p53 ssDNA]. Preliminary experiments revealed integral problems for the SMB assay as almost 100% of p53 bound to the beads in the absence of DNA using a recommended buffer of 10 mM Tris, pH 7.5, 50 mM KCl, 1 mM DTT (Figure 3.7 A). To reduce this unspecific binding, a separate incubation of free biotin molecules was introduced following bead-DNA incubation, with the aim of reducing exposed streptavidin on the bead surface (Figure 3.7 B). Although non-specific p53 binding to ‘No DNA’ was reduced, a significant amount was still interacting with the beads. To further reduce non-specific binding, BSA was introduced to the buffer used prior to addition of the protein of interest and during all wash steps (Figure 3.7 C). With the introduction of BSA, non-specific protein binding was significantly decreased to less than 50%.

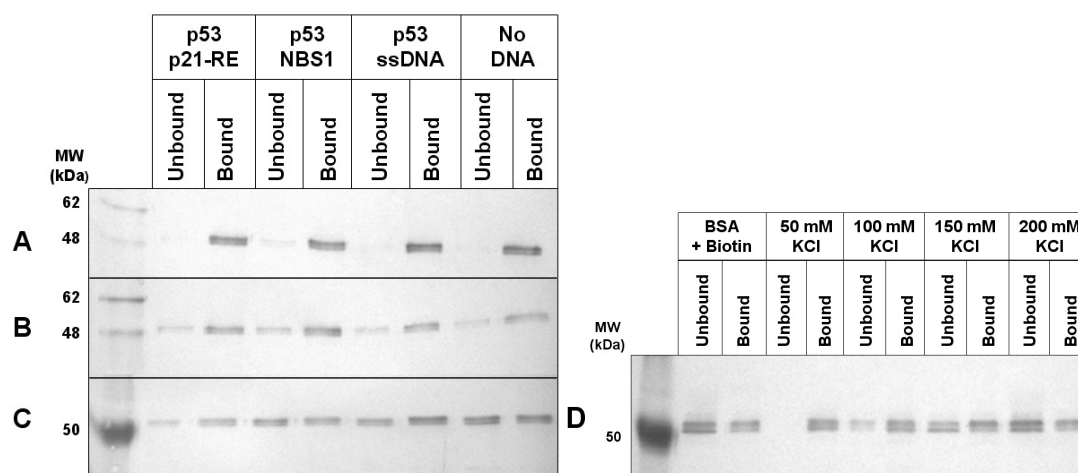


Figure 3.7 Reducing non-specific p53 binding to M-280 Streptavidin coated beads. Western blot results from assays looking at p53 binding dsDNA containing its consensus binding site ([p53 p21-RE]), non-specific dsDNA ([p53 NBS1]), non-specific ssDNA ([p53 ssDNA]) and to SMBs alone (No DNA). From each assay a sample of unbound and bound protein is collected. **A** - Results from an *in vitro* SMB assay with no optimisation of conditions. **B** - Results from an *in vitro* SMB assay with a free biotin incubation following DNA addition. **C** - Results from an *in vitro* SMB assay with addition of 400 µg / ml BSA incubation in all wash buffers. **D** - Results from an *in vitro* SMB assay with both biotin and BSA optimisations and effects of increasing [KCl] on non-specific p53 interactions with SMBs.

Varying ionic strength can have several implications in assays involving DNA and protein interactions. Increasing salt concentrations can facilitate DNA stabilisation and so oligonucleotides used during these assays were more likely to remain in their desired conformations. As electrostatic force strongly influences intermolecular interactions between the negatively charged DNA phosphate backbone and basic amino acid residues in proteins, cation concentration in a solution significantly contributes to the thermodynamic and kinetic properties of the specificity and stability of protein-DNA interactions (Oda and Nakamura, 2000). In general, binding affinity decreases as salt concentration increases and effectively, weaker non-specific protein-DNA interactions can be distinguished from stronger high-affinity interactions. It was anticipated that increasing the concentration of KCl would also decrease p53-SMB non-specific interactions as this phenomenon may also be affected by ionic strength. Figure 3.7 D shows results from an *in vitro* SMB assay investigating p53 binding to SMBs at various KCl concentrations. The results indicate p53-SMB interactions decrease as KCl increases, suggesting this interaction is ionic-strength dependent to an extent.

3.3.2 Dynabeads® Protein-G Analysis.

An alternative *in vitro* DNA binding assay used during this study employs Dynabeads® Protein-G beads (PGBs) to which immunoglobulins attach via their Fc part, enabling immunoprecipitation of proteins that may be bound to DNA. For *in vitro* PGB assays (Section 2.8) it is important to identify if the non-specific interactions associated with the SMBs were also a characteristic of these different beads. Initially, assays were performed independent of DNA in order to distinguish if p53 bound non-specifically to beads and to confirm that p53 (DO-1) would be an appropriate antibody for immunoprecipitation. The manufacturer (Invitrogen™) advises to use approximately 5 µg of antibody for every 50 µl of PGBs used. Previous studies on p53 have shown that 12.5 µl PGBs can be used to study 50 ng of p53 using 100 ng of p53 (DO-1) for immunoprecipitation (Fojta et al., unpublished data). As 100 ng is significantly less than the manufacturer's recommendation (1250 ng for 12.5 µl PGBs) it was important to confirm this amount was sufficient to immunoprecipitate 50 ng of p53 (Figure 3.8).

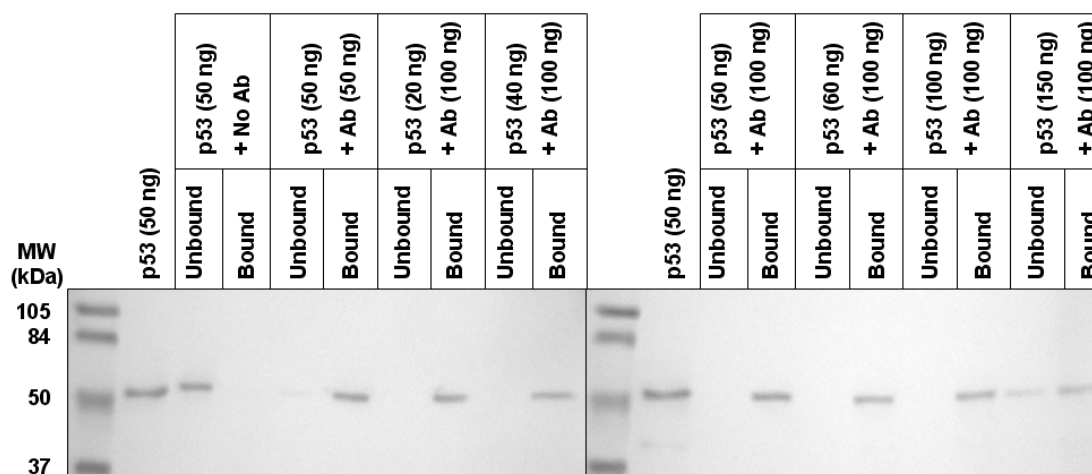


Figure 3.8 Western blot analysis of p53-PGB non-specific interactions and characterising relevant p53 (DO-1) antibody amounts. *In vitro* PGB assays were performed without the inclusion of DNA to identify if p53 bound non-specifically to PGBs. Various amounts of p53 and p53 (DO-1) antibody were also investigated to determine the appropriate amounts required for subsequent assays.

Contrary to M-280 Streptavidin coated beads, p53 did not appear to interact with protein-G beads, indicating the non-specific interactions previously observed (Figure 3.7) were likely to be due to the presence of streptavidin. Inclusion of an incubation of only 50 ng of p53 (DO-1) prior to addition of 50 ng p53 was enough for almost 100% of the p53 to attach to PGBs, and this was increased to 100% when 100 ng of p53 (DO-1) was present. It was also identified that 100 ng of p53 (DO-1) would allow 100 ng of p53 to precipitate, however this amount of p53 (DO-1) was not enough for attachment of 150 ng p53. It was accepted that for all following *in vitro* PGB assays that 100 ng of p53 (DO-1) would be appropriate to allow immunoprecipitation of 50 ng p53.

As the *in vitro* PGB assay analysed protein-DNA interactions by detection of bound DNA it was essential to confirm that PGBs with or without addition of antibodies would not precipitate DNA in the absence of protein. Figure 3.9 shows results from a PGB assay performed to investigate non-specific attachment of linear and supercoiled pUC18 to PGBs. The experiments were performed as described (Section 2.8) except no protein was included.

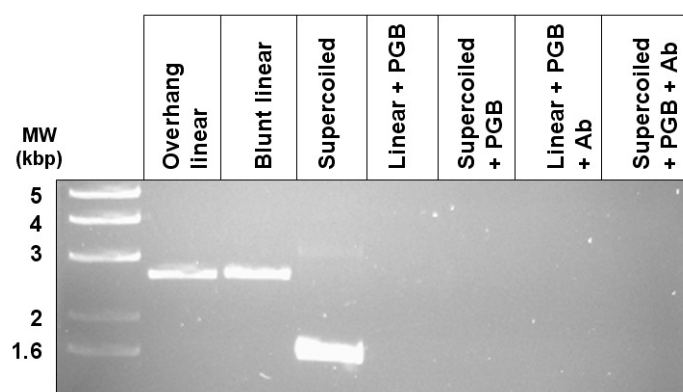


Figure 3.9 *In vitro* PGB assay results investigating PGB and pUC18 interactions. Linear pUC18 with overhang and blunt ends were analysed by electrophoresis on 1% agarose gels alongside supercoiled pUC18 which are shown in the first 3 lanes. The last 4 lanes are results from a PGB assay exclusive of protein with or without p53 (DO-1) antibody.

It was discovered that none of the pUC18 molecules studied demonstrated any interactions with PGBs or p53 (DO-1) antibody. These findings were promising as the *in vitro* PGB assay did not require further optimisation in order to prevent non-specific interactions using the assay conditions tested.

3.4 *In vitro* SMB DNA Binding Analysis of Human p53.

Over the last 20 years, several studies have attempted to elucidate the versatile and complex interactions p53 exerts on DNA that, as yet, remain only partially understood. It has been ascertained that p53-DNA interactions vary depending on both DNA sequence and structure, and also the status of the p53 protein itself (Kim and Deppert, 2006). It is apparent that cooperation between p53s CDBD and regulatory CTD ultimately determine the mode of interaction that occurs with DNAs consisting of specific consensus sequences, unusual non-B DNAs and even linear non-specific DNAs.

In vitro studies mainly executed using EMSA identified that, for initiation of sequence-specific DNA binding activity to occur, p53 required modification or complete removal of the CTD, which was at first thought to function as an inhibitor of consensus sequence binding. This led to the proposal of the ‘latency’ model based on the concept that unmodified p53 did not bind to consensus DNA. Recently, it was shown that the CTD actually facilitated consensus sequence binding (McKinney et al., 2004) by localising p53 to DNA and allowing for diffusion in reduced dimensions, effectively disproving the ‘latency’ model with respect to sequence-specific interactions. It was believed misleading data was obtained from EMSA assays investigating p53 binding to short linear DNA targets (Kim and Deppert, 2006), which

often required introduction of PAb421 antibody to differentiate between specific and non-specific p53-DNA interactions.

A previous study that employed EMSA to investigate p53 interactions with short DNAs (Walter et al., 2005) identified high structure-specific binding affinity of unmodified p53 to hairpin structures composed of CTG•CAG repeats compared to linear DNAs composed of a similar sequence. These results are consistent with unmodified p53 functioning as part of a global genomic surveillance system that could detect unstable DNA prone to detrimental rearrangements. However, for activation of p53 sequence-specific binding to linear consensus DNA from the Mdm2 promoter under the same experimental conditions, the introduction of PAb421 was required. As ‘latent’ p53 failed to bind to consensus DNA during these assays, the reliability of the detected interactions between p53 and unusual structures may just be an artefact of employing EMSA to detect p53 binding to short DNAs.

Determination of protein-DNA binding activity detected during SMB reactions initially requires the analysis of samples representing bound and unbound protein by SDS-PAGE and western blotting. Immunodetection with protein-specific antibodies then identifies the amounts of protein in each relevant lane that is proportional to the intensity of the observed bands. An example of how data is analysed following SMB assays is shown in Figure 3.10 (this figure is not experimental data and will not be described. DNA1 – DNA4 are hypothetical protein binding substrate DNAs).

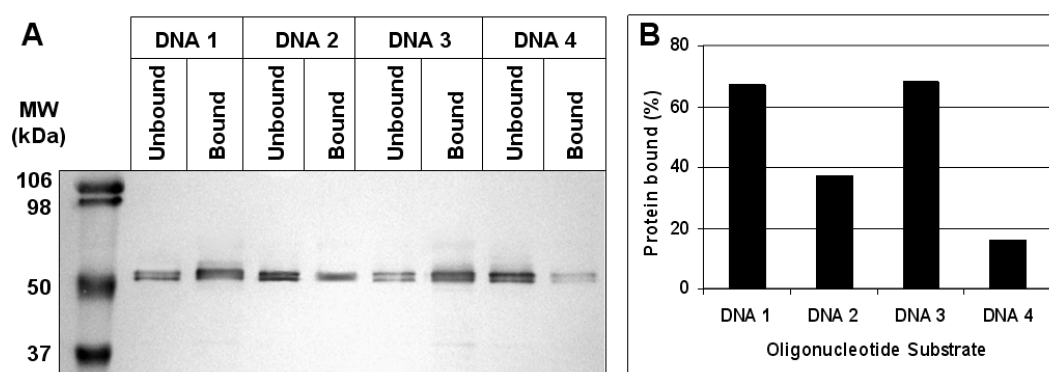


Figure 3.10 Example of the quantification of percentage binding by specific proteins during *in vitro* SMB DNA binding assays. For each DNA substrate investigated, samples representing bound and unbound protein are collected and analysed by SDS-PAGE and western blotting. Following immunodetection, band intensities are measured using ImageJ allowing for calculation of percentage binding. **A.** Western blot and immunodetection analysis of protein collected following an SMB assay. **B.** Data obtained following measurement of band intensities and conversion into values representing percentage binding.

Quantification of band intensities is performed using the program ImageJ (Figure 3.10) that allows for the identification of precise values for the amounts of protein bound or unbound. This method of analysis was performed during all *in vitro* SMB assays investigating human p53 binding to various oligonucleotide substrates.

3.4.1 How Ionic Strength Affects Human p53-DNA Interactions.

Protein-DNA interactions are often significantly dependent upon electrostatic forces in the surrounding environment and so, during *in vitro* protein-DNA binding analysis, it is important to relate observations to an estimate of the intracellular salt concentrations in which protein-DNA interactions occur naturally. Ionic strength impacts the stability of protein-DNA complexes as electrostatic attraction between the negatively charged phosphate backbone of DNA and positively charged residues on a protein's surface often determine the strength of an interaction. Electrostatic attraction is particularly important during the random molecular association that occurs during non-specific protein-DNA interactions (Jen-Jacobson, 1997). Increasing the ionic strength of the surrounding environment effectively neutralises electrostatic attractions and only specific protein-DNA interactions influenced by alternative thermodynamic determinants will be maintained. Although intracellular salt concentrations vary between organisms, investigations in *E. coli* determined the ionic environment to be equivalent to 160 mM K⁺ (Kao-Huang et al., 1977), a value that is thought to be a conservative estimation. Human cellular ionic values are believed to be similar (Becker et al., 2009).

It is not possible to precisely recreate conditions equivalent to the intracellular environment during *in vitro* DNA binding analysis, however as electrostatic interactions are so crucial during protein binding it would be appropriate to have comparable intracellular salt concentrations during *in vitro* assays. Previous EMSAs investigating p53-DNA interactions were frequently performed at salt concentrations considerably lower than the predicted intracellular values. For example, concentrations of 25 mM KCl (Espinosa and Emerson, 2001; McKinney and Prives, 2002), 50 mM KCl (Kim et al., 1997; Walter et al., 2005) or 50 mM NaCl (Gohler et al., 2005; Gohler et al., 2002) have been utilised. To examine how electrostatic forces govern p53-DNA interactions, SMB assays to determine p53 binding to various substrate DNAs at varying concentrations were performed. The substrate DNAs initially investigated included [p53 p21-RE] (a dsDNA containing the response element from the p21 promoter), [p53 NBS1] (a dsDNA containing a sequence from the p21 promoter that is distant from the response element) and [p53 ssDNA] (a ssDNA molecule with no sequence from any p53 response elements). As a control, reactions containing no DNA were also

analysed as non-specific p53-SMB interactions had been previously identified. Assays measuring p53-DNA interactions at 50 mM and 220 mM KCl provided results that displayed the highest differences in activity (Figure 3.11):

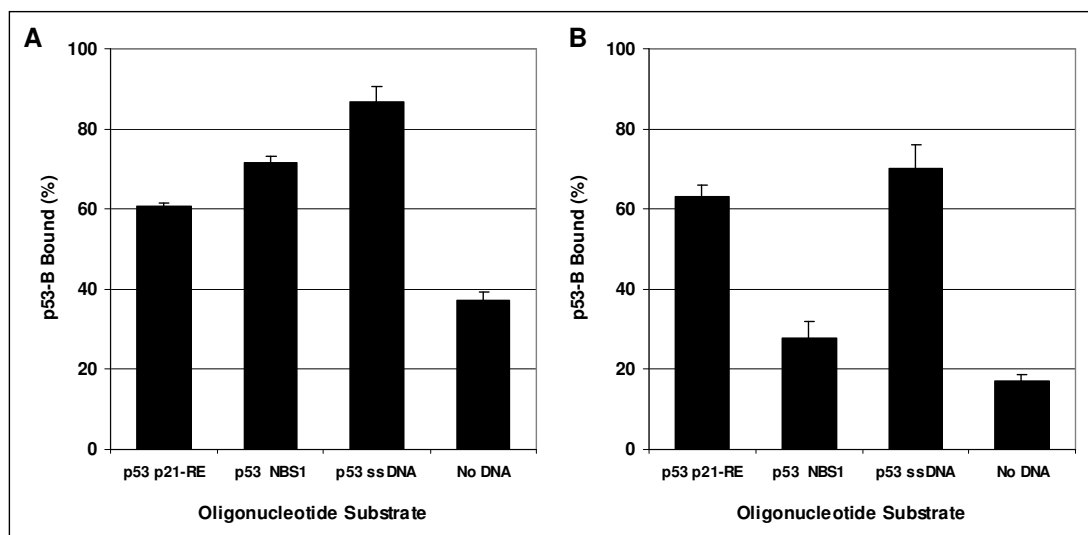


Figure 3.11 *In vitro* SMB analysis of human p53-B (over-expressed in bacterial cells) interactions with various DNAs at 50 mM and 220 mM KCl. Averages were taken from a minimum of 6 experiments. Standard errors are shown. **A.** Assays investigating human p53-B interactions with DNAs at 50 mM KCl. **B.** Assays investigating human p53-B interactions with DNAs at 220 mM KCl.

Table 3.1 Statistical analysis of results obtained during SMB assays investigating p53-B interactions with DNAs at 50 mM and 220 mM KCl. Two-tailed t-test analysis was performed on data obtained from assays shown in Figure 3.11. Percentage binding values for each DNA substrate from both assays at 50 mM and 220 mM were compared.

DNA substrate	<i>t</i> value	Degrees of freedom (of <i>t</i>)	Critical values of <i>t</i> (<i>p</i> = 0.05)	Binding difference
p53 p21-RE	0.3	15	2.13	Not significant
p53 NBS1	5.1	15	2.13	Significant
p53 ssDNA	1.5	6	2.45	Not significant
No DNA	5.3	7	2.37	Significant

The results shown in Figure 3.11 highlight the considerable influence that the ionic strength of reaction environments has upon p53-DNA interactions and how it is difficult to distinguish between non-specific and specific interactions at low salt concentrations. This is apparent when p53 binding to [p53 NBS1] is compared under both KCl concentrations where statistical analysis (Table 3.1) confirmed that there was significantly less p53 binding to this substrate at 220 mM KCl compared to at 50 mM. However, for both [p53 p21-RE] and [p53 ssDNA], there was no significant difference in binding between at different ionic strengths. As p53 should have no specific affinity towards [p53 NBS1], it is likely the interactions seen to this

substrate are non-specific, and so are heavily dependent on electrostatic interactions. However, as p53 binding to both [p53 p21-RE] and [p53 ssDNA] remained relatively high even at 220 mM KCl, it is expected that the observed binding is specific and, therefore, thermodynamically more stable. Although p53-interactions towards these 2 DNAs may appear similar, the mode of binding is likely to be different as binding to [p53 p21 RE] is believed to be largely sequence-specific whereas binding to [p53 ssDNA] would be determined by structure. This experiment has illustrated that p53 can distinguish between non-specific dsDNA and sequence-specific dsDNA without the requirement of modifications to the CTD and, therefore, it does not support previous 'latency' models (Anderson et al., 1997; Hupp and Lane, 1994). Data from these assays concerning non-specific p53 interactions with SMBs agreed strongly with previous results (Section 3.3.1.2) in that at increased ionic strength, p53 association with SMBs was significantly reduced. It was, therefore, concluded from these assays that salt concentrations had a considerable effect on p53 activity and, to allow discrimination between specific and non-specific interactions, all future p53 SMB assays would be performed at 220 mM KCl.

3.4.2 Comparing Human p53s Derived From Eukaryotic and Prokaryotic Host cells.

Speculation regarding the requirement for post-translational modifications for the induction of sequence-specific DNA binding by p53 has been a consistent feature of several investigations into p53-DNA interactions. It was of interest to elucidate if binding activity between human p53 proteins derived from either eukaryotic or prokaryotic host cells exhibited varied DNA binding activity due to limitations of prokaryotic post-translational modifications compared to eukaryotes. *In vitro* SMB assays were performed on 2 separate human p53 proteins that had been over-expressed in either insect (p53-I) or bacterial (p53-B) cell lines. For the determination of precise binding activities, several DNA substrates were employed including [p53 p21 RE], [p53 NBS1] and [p53 ssDNA] that had been previously studied during ionic strength assays and, additionally, dsDNA molecules that contained internal hairpins of various sizes consisting of CTG repeats were also analysed. It was of interest to determine if binding by p53 to CTG-hairpins that had been reported in a previous study (Walter et al., 2005) was indeed a specific interaction or an artefact resulting from EMSA analysis. As interactions towards DNAs containing internal hairpins is likely to be a structural-specific interaction, to ascertain the dependence of the interaction on size of the hairpin structure, hairpins consisting of 1 CTG repeat ([p53 (CTG)₁ HP]), 3 repeats ([p53 (CTG)₃ HP]) or 7 repeats ([p53 (CTG)₇ HP]) were investigated.

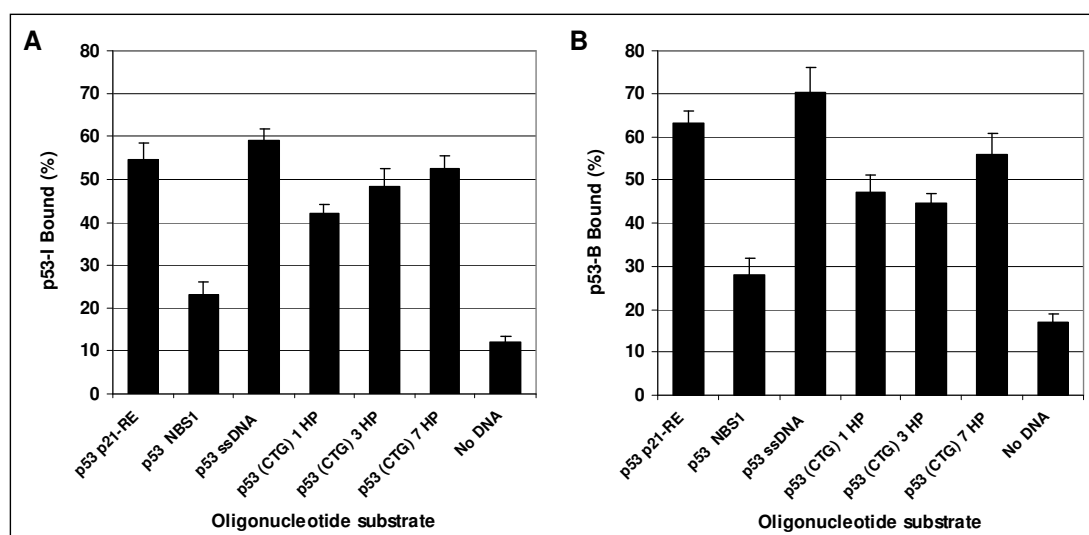


Figure 3.12 *In vitro* SMB analysis of human p53-I and p53-B interactions with various DNAs at 220 mM KCl. Averages were taken from a minimum of 6 experiments. Standard errors are shown. **A.** Assays investigating human p53-I interactions with DNAs. **B.** Assays investigating human p53-B interactions with DNAs.

Table 3.2 Statistical analysis of results obtained during SMB assays investigating human p53-I (normal font) and p53-B (*italics*) interactions with various DNAs at 220 mM KCl. Two-tailed t-test analysis was performed on data obtained from assays shown in Figure 3.12. Percentage binding values of both p53 proteins towards [p53 p21 RE] were compared to values representing binding to alternative DNAs. The differences between non-specific DNA binding and non-specific SMB binding were also compared.

DNA substrates compared	<i>t</i> value	Degrees of freedom (of <i>t</i>)	Critical values of <i>t</i> (<i>p</i> = 0.05)	Binding difference
p53 p21-RE v p53 NBS1	6.3	8	2.31	Significant
p53 p21-RE v p53 ssDNA	0.9	19	2.09	Not significant
p53 p21-RE v p53 (CTG) ₁ HP	2.5	7	2.37	Significant
p53 p21-RE v p53 (CTG) ₃ HP	1.3	7	2.37	Not significant
p53 p21-RE v p53 (CTG) ₇ HP	0.4	8	2.31	Not significant
p53 NBS1 v No DNA	3.1	7	2.37	Significant
<i>p53 p21-RE v p53 NBS1</i>	8.3	28	2.05	<i>Significant</i>
<i>p53 p21-RE v p53 ssDNA</i>	1.3	19	2.09	<i>Not significant</i>
<i>p53 p21-RE v p53 (CTG)₁ HP</i>	2.1	20	2.09	<i>Significant</i>
<i>p53 p21-RE v p53 (CTG)₃ HP</i>	2.6	20	2.09	<i>Significant</i>
<i>p53 p21-RE v p53 (CTG)₇ HP</i>	1.5	24	2.06	<i>Not significant</i>
<i>p53 NBS1 v No DNA</i>	2.4	20	2.09	<i>Significant</i>

It can be concluded from the data shown in Figure 3.12 that both p53-I and p53-B exhibit almost indistinguishable DNA binding properties. As both proteins should be identical in amino acid sequence, excluding the polyhistidine tag present on p53-B, the main factor determining variations in binding would have been thought to arise as a consequence of alternative over-expression systems. As binding is very similar, the results indicate that any post-translation modifications present on p53-I and not on p53-B did not influence interactions with DNA. Statistical analysis (Table 3.2) revealed that, for both p53s, a significant difference in binding was observed between [p53 p21 RE] and [p53 NBS1], however, both p53s did not show significant differences in binding between [p53 p21 RE] and [p53 ssDNA]. As the ssDNA substrate did not contain any response element sequence, this interaction is a structure-dependent interaction that is a recognised activity of p53. Interestingly, the data obtained suggests the interaction of p53 with ssDNA is similar to that observed with consensus dsDNA, although the naturally occurring topology of the p21 promoter, which has not been accounted for in this assay, may induce a tighter interaction.

One of the primary aims of the SMB assay was to identify if DNAs containing internal hairpins consisting of CTG repeats would induce p53 binding. As p53 has natural non-specific affinity towards linear dsDNA it was assumed that for an interaction with hairpin structures to be a specific interaction, then p53 binding would need to be greater than that observed to linear dsDNA of similar size. Interestingly, it was recognised that the introduction of a single CTG repeat induced considerable p53 binding compared to non-specific linear dsDNA, however binding remained significantly lower than sequence-specific binding. The observed impact of the presence of a single trinucleotide repeat may at first appear cryptic, but when p53's recognised ability to associate with mismatches in DNA is considered (Degtyareva et al., 2001), the finding that a 3-base bulge induces increased binding seems less remarkable. The only difference observed in binding activity between p53-I and p53-B was towards [p53 (CTG)₃ HP] as p53-I binding was not significantly different compared to [p53 p21 RE] binding, however p53-B was. This difference seems of less importance when binding to [p53 (CTG)₇ HP] is compared, as both p53s interact comparatively well, and are statistically not significantly different than consensus DNA binding. This result implies p53-B binding to [p53 (CTG)₃ HP] is an anomaly in that binding was unusually low in the experiments performed. This indication is reinforced further as it appears that p53 binding to DNAs containing internal hairpins generally increases as the hairpins increase in size, and it is only binding of p53-B to [p53 (CTG)₃ HP] that does not correspond to this consensus. As p53 binding to hairpins consisting of 7 CTG repeats is not significantly different compared to consensus DNA binding, these results are comparable with previous findings (Walter et al., 2005) and imply similar regions that arise naturally in

genomes may induce p53 interactions. The results shown in (Figure 3.12) also prove that, for both p53s, the non-specific interaction towards linear dsDNA is a significant result when compared to binding to ‘no DNA’. This data agrees with the currently accepted model that non-modified p53 interacts with linear dsDNA independent of sequence, a function believed to facilitate location of specific regions in DNA.

3.4.3 Human p53-Interactions with Unusual DNAs.

It was believed previous SMB assays had detected at least 3 modes of p53 binding: sequence-specific, structure-specific and non-specific DNA interactions. The considerable difference in p53 binding to linear dsDNA containing or not containing a response element indicated that although p53 does show binding affinity towards dsDNA of any sequence, the association that occurs with sequence-specific DNA is noticeably tighter. The observed binding to ssDNA and to dsDNA containing internal hairpins was expected to be largely sequence-independent, however further clarification regarding p53 interactions with unusual DNAs was required. SMB assays were carried out using dsDNAs consisting of various unusual features, ranging from single mismatches to DNAs with loop regions. To further determine if p53 interactions towards internal hairpin molecules were solely based on structure, assays were also performed investigating p53 binding to dsDNA containing internal hairpins consisting of different TRS. For all subsequent assays, p53-B was used, but given the data reported above, it was expected similar observations would be made for p53-I.

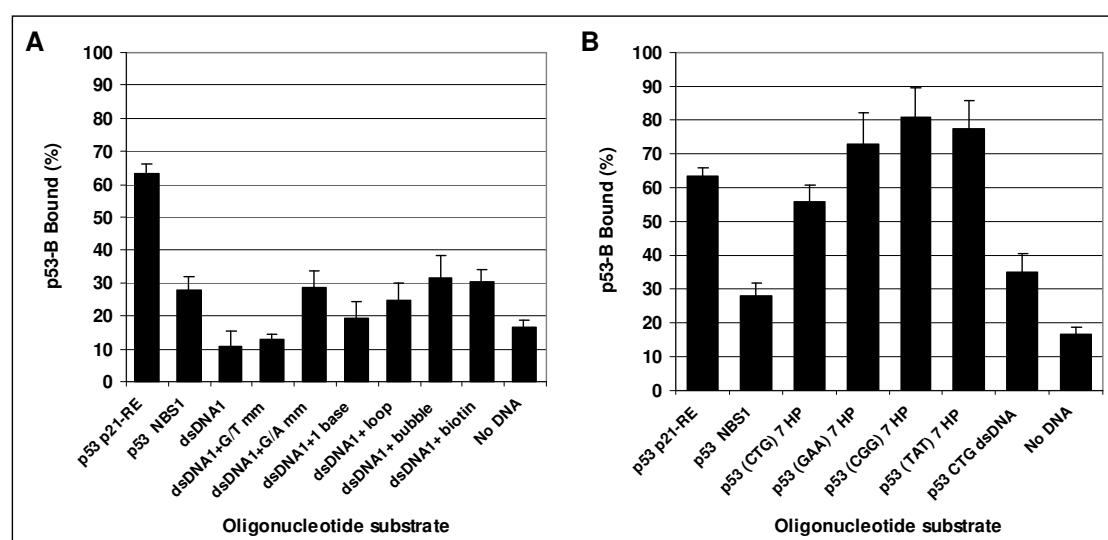


Figure 3.13 *In vitro* SMB analysis of human p53-B interactions with dsDNA containing unusual features at 220 mM KCl. Averages were taken from a minimum of 3 experiments. Standard errors are shown. **A.** Assays investigating human p53 interactions with dsDNAs containing unusual features. **B.** Assays investigating human p53 interactions with DNAs containing internal hairpins consisting of varied TRS.

Table 3.3 Statistical analysis of results obtained during SMB assays investigating human p53-B interactions with DNAs containing internal TRS hairpins. Two-tailed t-test analysis was performed on data obtained from assays shown in Figure 3.13 B. Percentage binding values of p53 towards [p53 (CTG)₇ HP] were compared to values representing binding to alternative hairpin DNAs or linear dsDNA containing CTG repeats. p53 binding to linear CTG repeats was also compared to binding to non-specific dsDNA.

DNA substrate	<i>t</i> value	Degrees of freedom (of <i>t</i>)	Critical values of <i>t</i> (<i>p</i> = 0.05)	Binding difference
p53 (CTG) ₇ HP v p53 (GAA) ₇ HP	1.9	12	2.18	Not significant
p53 (CTG) ₇ HP v p53 (CGG) ₇ HP	2.8	12	2.18	Significant
p53 (CTG) ₇ HP v p53 (TAT) ₇ HP	2.5	12	2.18	Significant
p53 (CTG) ₇ HP v p53 CTG dsDNA	2.5	12	2.18	Significant
p53 CTG dsDNA v p53 NBS1	1.0	16	2.12	Not Significant

The results shown in Figure 3.13 A indicate that p53 binding to the various substrate DNAs was relatively low compared to structural-dependent binding observed previously. Surprisingly, it was apparent that p53 had a preference towards [p53 NBS1] over [dsDNA1], an unexpected result as both contain non-sequence specific DNA. The inclusion of a G/T mismatch in the latter led to slightly increased binding, however when a G/A mismatch was present, p53 binding was visibly enhanced and even greater than [p53 NBS1] binding. This data correlates well with previous studies (Degtyareva et al., 2001) that found p53 bound with higher affinity towards A/G mismatches and with lower affinity towards G/T mismatches. One possibility for the preference of A/G mismatches over G/T is that the additional imidazole ring present when 2 purines are paired may have a more substantial impact on the DNA structure. The inclusion of a single base insert or a loop region also increased p53 binding, however not to the extent of the G/A mismatch substrate. As the inclusion of either a base insert or loop structure would induce conformational changes within the DNA, p53 binding may have been anticipated to have been higher. A series of mismatches contained within the dsDNA ([dsDNA1+bubble]) also increased binding, but, again, not as extensively as expected. Addition of biotin to both ends of the dsDNA [dsDNA1+biotin] also appeared to induce interactions with p53, however the reason for this is not clear. As a biotin molecule is present at both ends of this dsDNA, it is likely it will be juxtaposed closely to the SMB, which would be expected to make protein binding more difficult. However, as both ends of this molecule are fixed, the prospect of twisting of this DNA is possible, a feature that has been shown to increase p53 binding (Gohler et al., 2002; Palecek et al., 2001).

Although p53 does appear to demonstrate binding preference towards DNAs containing features likely to induce unusual conformations on the DNA, p53 binding shown in Figure 3.13 A to the various structural derivatives of [dsDNA1] was generally low. This is apparent even when [dsDNA1] was used for p53 binding substrate as it was expected p53 binding would be similar as that identified when [p53 NBS1] was substrate. Apart from sequence, the only difference between these 2 substrates was that [dsDNA1] was 7 base pairs shorter than [p53 NBS1], although, at 30 base pairs, this should still be large enough to accommodate p53. However, if p53 has a binding preference for larger DNAs, this would explain why binding to [dsDNA1] and its derivatives was relatively low.

SMB assays were also performed that analysed p53 interactions with dsDNAs containing internal hairpins consisting of 7 repeats of various TRS sequences (Figure 3.13 B). It was important to determine if the interactions observed towards CTG hairpins was structure-dependent or if sequence also influenced p53 affinity. It was found that p53 bound relatively tightly to all hairpin DNAs investigated, though interestingly, sequence did appear to influence p53-DNA binding. Statistical analysis (Table 3.3) revealed that there was no significant difference between p53 binding to [p53 (CTG)₇ HP] and [p53 (GAA)₇ HP], however, significant differences were observed between [p53 (CTG)₇ HP] and both [p53 (CGG)₇ HP] and [p53 (TAT)₇ HP]. As all DNAs were of identical size, the differences observed were sequence-dependent. However, it is likely that the different sequences induced varying conformations on the hairpins and so, ultimately, it could be argued that the overall structure determined interaction strength. Hairpins present in both [p53 (CGG)₇ HP] and [p53 (TAT)₇ HP] would be very different since the CGG hairpin would be both bulky as a consequence of the abundance of guanine nucleotides, but also tightly associated because of the strong cytosine and guanine base pairing. Contrastingly, the TAT hairpin may be predominantly single-stranded due to the misalignment of adenine and thymine nucleotides.

Although the data indicates sequence does influence p53-hairpin interactions, it is apparent that binding is still relatively high irrespective of the nucleotide constituents. The hypothesis that binding to hairpins is heavily structural dependent is confirmed by comparing [p53 (CTG)₇ HP] and [p53 CTG dsDNA]. Whereas both DNAs are identical in size and sequence, only [p53 (CTG)₇ HP] contains an internal hairpin and, therefore, any differences in p53 binding would effectively be structure-dependent. As shown in Figure 3.13 B, p53 binding was significantly greater to [p53 (CTG)₇ HP], reinforcing the notion that structure is the important determinant in p53 binding to TRS.

3.5 Chromatin Immunoprecipitation.

Chromatin immunoprecipitation (ChIP) is an experimental technique that can detect protein interactions with specific genomic regions *in vivo*. This method relies on the formation of a covalent crosslink between the protein of interest and any region of DNA that the protein is bound to. Formaldehyde is commonly used for this process as it reacts with DNA bases and primary amines on amino acids to create a covalent crosslink between both molecules. Cells are lysed following cross-linking and extracts are sonicated to shear chromatin into 0.2 – 1 kb fragments. An antibody specific for the protein of interest is then used to immunoprecipitate protein-DNA complexes from which DNA is released by reversing cross-links with heat, and remaining proteins are removed with proteinase K treatment. Techniques such as phenol:chloroform extraction can then be used to clean precipitated DNA, which can then be subjected to PCR to identify the presence of specific genomic regions if predicted binding sites are known. This procedure is described in more detail in Section 2.13.

In vitro human p53-DNA binding assays had indicated that genomic regions prone to conforming unusual secondary structures may behave as natural p53 binding sites. In order to elucidate this activity further it was of interest to characterise this interaction *in vivo* using ChIP. Prior to investigating interactions between p53 and unstable genomic regions, it was decided to perform ChIP assays that would detect p53 binding to known consensus sites. During *in vitro* SMB assays, a region from the p21 promoter was used as a substrate representing a p53 consensus binding site, therefore it was decided to investigate binding to this region *in vivo*. The Mdm2 promoter was also investigated during ChIP assays as it has been identified that p53 has strong affinity towards this region (Weinberg et al., 2005). As a negative control, a region from the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoter was chosen as this gene is often constitutively expressed in most cells and is not believed to be a downstream target gene of p53. A previous ChIP study investigating *in vivo* p53-DNA interactions (Kaesler and Iggo, 2002) also explored p53 binding to these 3 genomic regions and, for consistency, this current study utilised identical primers during PCR analysis.

ChIP assays had previously been performed at the University of East Anglia (UEA) to detect *in vivo* transcription factor Sp1 DNA binding in HeLa cells (an immortal cell line derived from human cervical cancer cells) (personal communication). Regardless of the specific degradation of p53 by HPV in cervical cancer cell lines, a transcriptionally active p53 that demonstrates sequence-specific DNA binding activity has previously been detected (Hoppe-Seyler and Butz, 1993), but the specific transcriptional activity of endogenous p53 in viral-induced cancers remains poorly understood. The same study investigated the functional

status of endogenous p53 in HeLa cells infected with HPV18 and detected a low amount of p53 that possessed sequence-specific DNA binding activity that was detected using gel retardation assays. Addition of exogenous wild-type p53 significantly increased p53-mediated transcription, which indicated that the limited amount of endogenous p53 in HeLa cells is not enough to induce p53-mediated responses. It was postulated that tumourigenic phenotypes of cervical cancers do not require the complete inactivation of p53, but that a threshold exists at which p53 is unable to exert its function. Therefore, in HPV infected cells such as HeLa, not all intracellular p53 is functionally inactivated and that remaining endogenous amounts varied between cells depending on viral protein expression.

3.5.1 Chromatin Analysis.

A crucial step during ChIP is the fragmentation of chromatin into 0.2 – 1 kb products. Unsuccessful fragmentation can result in the precipitation of large DNA molecules that contain unspecific regions, a factor that can significantly increase background or increase the possibility of false-positive results. Shearing of chromatin was performed by sonication, initially 10 x 30 seconds for each chromatin preparation. Results from these experiments revealed that although the majority of DNA was at approximately 0.5 – 2 kb, some fragments remained that were over 3 kb. The sonication procedure was optimised by reducing the opportunity for sample ‘foaming’ and the addition of 2 extra rounds of sonicating to all samples (Figure 3.14). Chromatin fragmentation analysis was encouraging as it was apparent that the fragments resulting from sonication had been successfully reduced in size. The smaller fragments would reduce precipitation of unspecific DNA and increase the reliability of ChIP results.

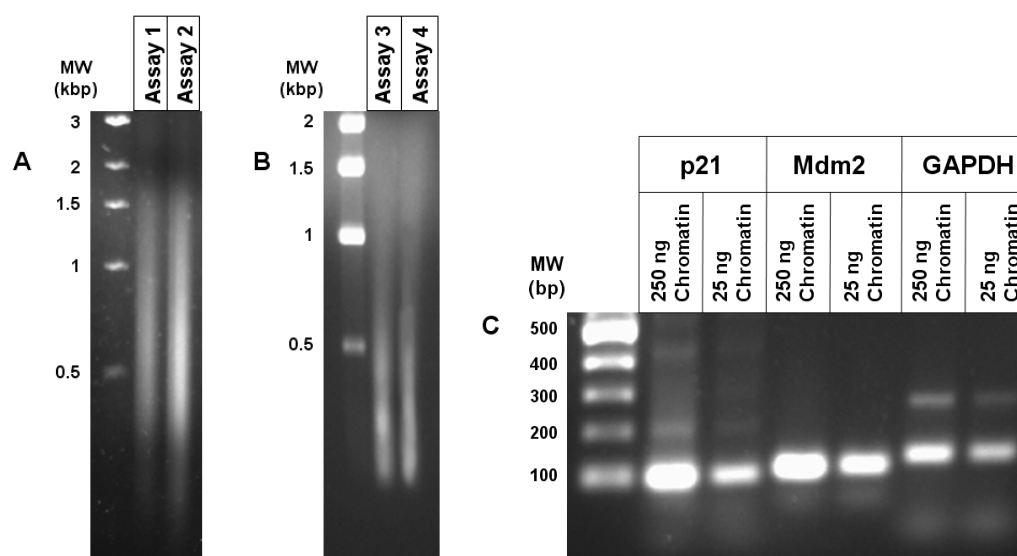


Figure 3.14 Electrophoretic analysis of chromatin fragments following shearing by sonication and testing for specific promoter regions by PCR. **A.** – Chromatin fragments prepared for 2 ChIP assays following 10 x 30 seconds of sonication. **B.** – Chromatin fragments prepared for 2 ChIP assays following 12 x 30 seconds of sonication. **C.** – PCR analysis of fragmented chromatin using primers designed to amplify p21, Mdm2 and GAPDH promoter regions using preliminary conditions. Predicted promoter regions sizes are: p21 – 105 bp, Mdm2 – 117 bp and GAPDH – 135 bp.

Chromatin was also analysed by PCR to confirm the genomic regions under investigation were present in the chromatin prior to ChIP. This also provided an opportunity to test primers and to identify appropriate conditions for ensuing PCRs. Approximately 25 ng and 250 ng of fragmented chromatin were analysed using primers specific for p21, Mdm2 and GAPDH promoter regions (Figure 3.14 C). The results indicated that the chromatin did contain all 3 promoters and that primers were working efficiently as bands were present in the expected locations. Irrelevant background products were low, however an extra band was apparent for GAPDH primers. To amend this, the annealing temperature for GAPDH primers were altered to favour the production of the specific fragment so that PCR conditions were identical to a previous study amplifying these promoter regions (Kaesler and Iggo, 2002).

3.5.2 *In vivo* p53-DNA Interactions.

ChIP assays were performed as described in Section 2.13.2. 3 samples of chromatin were collected from each assay; the first was initial chromatin input, which was a sample of fragmented chromatin collected following sonication; the second was chromatin immunoprecipitated using p53 (DO-1) antibody specific for p53; the third was chromatin that was immunoprecipitated using an irrelevant mouse monoclonal antibody. All 3 chromatin samples were analysed by PCR using primers designed to amplify regions of p21, Mdm2 and

GAPDH promoters. Due to the significant difference in the amount of DNA in the chromatin input sample compared to immunoprecipitated chromatin, the input chromatin was diluted 3-fold to decrease the differences in band intensity following electrophoresis and to allow more accurate analysis. Examples of PCR analysis from 3 ChIP assays from HeLa cells not irradiated with UV light are shown in Figure 3.15.

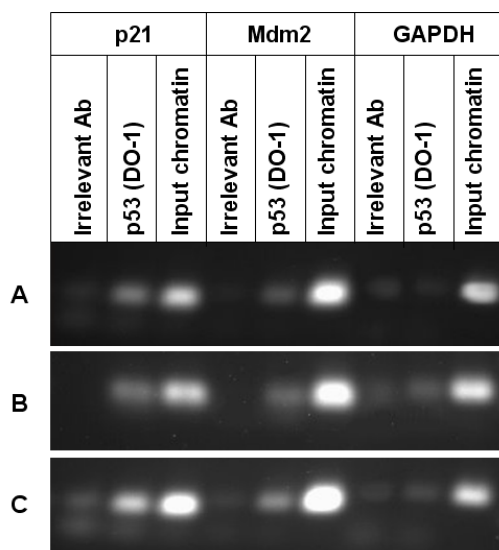


Figure 3.15 PCR analysis of 3 ChIP assays. Chromatin collected following sonication (input chromatin) was analysed by PCR alongside chromatin immunoprecipitated using an antibody specific for p53 (DO-1) and an irrelevant antibody. Primers designed to amplify regions of p21, Mdm2 and GAPDH promoters were used to detect the amounts of these regions in each sample. The amount of input chromatin used as a PCR template was diluted 3-fold to reduce differences in band intensities. **A**, **B**, and **C** represent 3 independent assays.

As Figure 3.15 demonstrates, the PCR products that formed depended on the initial amount of template DNA. For input chromatin samples, relatively intense bands were produced however, reactions that depended on immunoprecipitated chromatin produced less amplified DNA. An important factor was that chromatin immunoprecipitated using the irrelevant antibody contained very low amounts of the promoter regions studied. This was encouraging as it indicated that chromatin immunoprecipitated by p53 (DO-1) was determined by a specific interaction between p53 (DO-1) and p53. Further ChIP assays were performed, including investigations into p53-DNA in HeLa cells that had been irradiated with 20 J / m² UV light alongside cells not irradiated. PCR was used to detect immunoprecipitated chromatin.

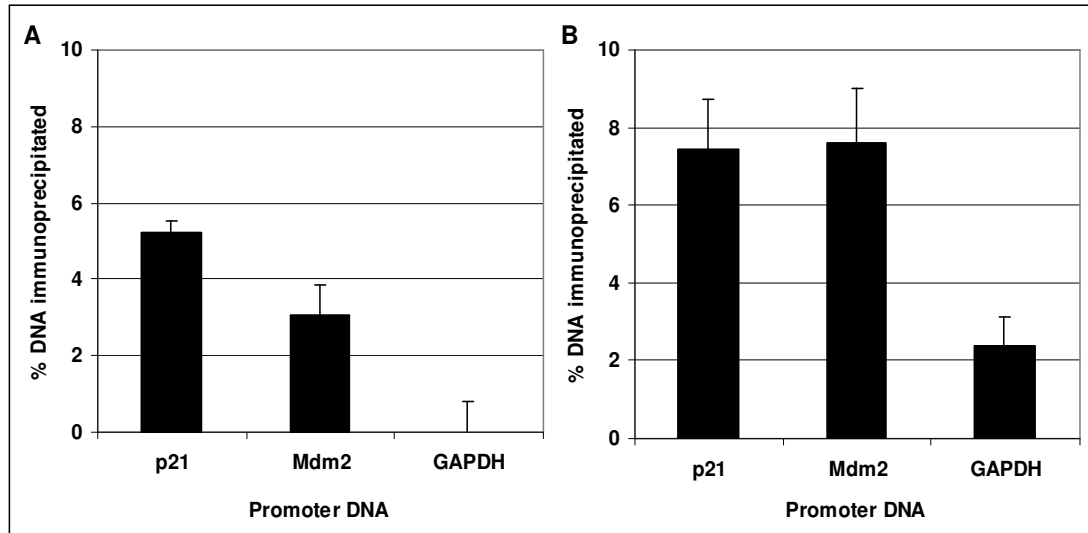


Figure 3.16 Summary of ChIP results investigating p53-DNA binding to p21, Mdm2 and GAPDH promoter regions in non UV-irradiated and UV-irradiated HeLa cells. Following ChIP assays, chromatin input, chromatin immunoprecipitated using p53 (DO-1) and chromatin immunoprecipitated using an irrelevant antibody were analysed by PCR and electrophoresis on 1% gels. Resulting band intensities were measured using ImageJ and percentages were calculated by comparing chromatin precipitated with p53 (DO-1) with chromatin input. Averages were taken from 3 ChIP assays. Standard errors are shown. **A.** ChIP results analysing p53-DNA binding in non UV-irradiated HeLa cells. **B.** ChIP results analysing p53-DNA binding in UV-irradiated HeLa cells.

A significant difference in binding was observed between p21 and GAPDH promoter regions ($t = 6.8$ ($p < 0.05$)), and also between the Mdm2 and GAPDH promoter regions ($t = 3.5$ ($p < 0.05$)) in HeLa cells not UV-irradiated (Figure 3.16 A). Therefore, results from ChIP assays indicate the existence of an active p53 containing the ability to bind sequence-specifically to DNA in HeLa cells, agreeing with previous *in vitro* studies (Hoppe-Seyler and Butz, 1993). As both p21 and Mdm2 are recognised p53-DNA binding elements the results suggest a clear preference for these regions compared to non-specific DNA. No significant difference was observed between p21 and Mdm2 regions ($t = 2.2$ ($p > 0.05$)).

UV-irradiation appeared to influence p53-DNA to an extent, however the pattern of binding was similar to non UV-irradiated HeLa ChIP assays (Figure 3.16 B). Again, p21 and Mdm2 had significantly higher p53 binding compared to the GAPDH promoter region ($t = 3.4$ ($p < 0.05$) and $t = 3.7$ ($p < 0.05$) respectively), and no difference was observed between p53 binding to p21 and Mdm2 regions ($t = 0.09$ ($p > 0.05$)). In comparison to non UV-irradiated HeLa cells, all 3 promoter regions had significantly higher binding (p21 regions: $t = 2.5$ ($p < 0.05$); Mdm2 regions: $t = 2.8$ ($p < 0.05$); GAPDH regions: $t = 2.3$ ($p < 0.05$)). As the increase in p53 binding to all 3 regions was very similar, it cannot be concluded that p53 sequence-specific DNA binding is increased following UV-irradiation, however it appears that p53-

DNA interactions in general increased in these cells. The chromatin immunoprecipitated containing the GAPDH promoter is likely to result from non-specific p53-DNA interactions via the p53 C-terminal domain, or alternatively p53 interactions with chromatin proteins. This non-specific chromatin may also account for a proportion of the chromatin in both p21 and Mdm2 assays. It was observed, however, that Mdm2 increased by the highest proportion following UV-irradiation as the percent chromatin bound more than doubled. These results show correlation with previous ChIP studies (Kaesler and Iggo, 2002) as p53 binding to p21 and Mdm2 promoters was significantly higher than GAPDH promoter binding, however the difference observed in this previous study was comparatively higher.

The percentage of chromatin immunoprecipitated during the ChIP assays was relatively high considering endogenous p53 in HeLa cells is recognised as being poorly expressed. Background non-specific p53-DNA binding may account for a proportion of immunoprecipitated chromatin as the p53 C-terminal domain will induce interactions in a sequence-independent mode. The high sensitivity of p53 (DO-1) antibody may also influence background binding as unspecific-protein interactions may arise if used at inappropriate concentrations. Alternatively, the sensitivity of p53 (DO-1) may be one reason why such low levels of endogenous p53 were able to be detected and so may have been crucial for these assays.

3.6 Discussion.

The activities of human p53 require tight coordination to ensure its inhibitory effects on cell growth are only operational when necessary. Post-translational modifications to specific sites on the p53 protein are accepted to be the principal mechanisms that govern elements including stability, cellular localisation and conformational changes (Woods and Vousden, 2001). The role of modifications including ubiquitination and phosphorylations on stability of p53 have been well characterised, however any direct influence that post-translational modifications have on the DNA binding properties of p53 remain to be fully understood. To this extent, 2 human p53 proteins that had been over-expressed in eukaryotic (p53-I) or prokaryotic (p53-B) host cells were investigated in this study to investigate if DNA binding activities differed as a consequence of production. It was believed that as prokaryotes do not contain advanced cellular machinery implicated in the protein modifications and folding required for some eukaryotic proteins, any differences observed between both p53 types could have resulted from differences in modifications or protein folding. Electrophoretic analysis of p53-I and p53-B revealed considerable differences in migration as p53-B appeared several kDa greater in size compared to p53-I. As both p53s were full-length wild-type versions of

human p53, this difference was not anticipated. Although p53-B was a recombinant version that had an extra 14 amino acids comprising the polyhistidine tag, these additional amino acids did not justify the difference observed. It is possible that the p53s analysed may have been alternate isoforms of p53 that arise following splicing of the human *p53* gene into splice variants prior to cloning into the respective hosts.

Prior to *in vitro* DNA binding analysis of both p53 types, the production of relevant oligonucleotide substrates was required. For all dsDNA molecules, appropriate ssDNAs were annealed and then analysed by native gel electrophoresis and subsequent chemiluminescent detection of biotin tags. Initially, annealing of substrates was inefficient, particularly to DNAs containing substantial secondary structures, and large amounts of ssDNA were present. This indicated that either dsDNA were unstable, and therefore denaturing, or that the annealing process was inefficient. Annealing conditions were optimised by increasing the ionic strength of the solution environment to neutralise the repelling effect of the negatively charged DNA backbones and promote association of complementing strands. Increasing the ratio of biotinylated: unbiotinylated ssDNAs to 1:4 respectively greatly improved incorporation of biotinylated ssDNA into dsDNA, resulting in excess unbiotinylated ssDNA that would be removed during SMB assay wash procedures. In some instances, analysis of dsDNAs containing internal hairpin molecules consisting of 7 TRS displayed multiple species of dsDNAs, which were accepted to be conformational isomers of the same DNA molecule that were also detected in a previous study (Walter et al., 2005). All DNAs that were employed as binding substrates during *in vitro* DNA binding assays were subject to structural analysis prior to use.

It was also crucial to analyse components of DNA binding assays in advance to identify factors that could ultimately influence the acquired data. DNAs that were both biotinylated and fluorescently tagged were used to measure the DNA attachment capacity of SMBs. It was found that both ssDNA and dsDNA attached to SMBs to a similar extent, indicating that any hindrance caused by the larger dsDNAs did not significantly reduce attachment. It was apparent however, that in order to attain the expected amount of DNA attached to SMBs, approximately 3 times of the manufacturers stated capacity of DNA should be incubated with beads. Analysis of DNA loss following washing procedures during assays suggested that very little ssDNA was lost, however it was apparent that a small amount of dsDNA was removed as indicated by a loss in fluorescence emission. As the interaction between biotin and streptavidin is one of the strongest non-covalent interactions known ($K_d \sim 4 \times 10^{-14}$ M), this slight loss may be explained by factors other than breakage of biotin-streptavidin bonds, such as DNA denaturing and the ensuing loss of the fluorescently labelled strand.

Non-specific protein interactions towards bead components of *in vitro* DNA binding assays would be highly disadvantageous as protein would be precipitated in conjunction with protein-DNA complexes independent of DNA, thereby creating false-positive results. Quantification of non-specific interactions towards SMBs was achieved by performing assays without the addition of DNA so that any protein precipitated would have been dependent on beads. Under recommended conditions, non-specific binding was considerable as almost all human p53 attached to SMBs without DNA present. To counteract this problem, an additional free biotin incubation step was introduced into the procedure in an attempt to reduce exposed streptavidin on the bead surface. The introduction of BSA in assay buffers was also investigated as it was assumed BSA may also interact non-specifically to the bead surface and, therefore, help to reduce bead surface exposure for proteins. Although beneficial, the most significant reduction in non-specific p53-SMB interactions was the increase in KCl concentration, indicating ionic strength was an important component of the observed interaction. It was also believed that increases in ionic strength of reaction environments would enhance double-stranded oligonucleotide stability and help prevent denaturing by neutralising negative charges of complementary strands. In contrast to SMBs, PGBs displayed no non-specific interactions with human p53 and specific antibodies were needed for protein precipitation. Non-specific association of DNA to PGBs was also examined with or without antibodies and it was determined that DNA did not associate with PGBs in either situation. Capacity analysis was also performed on PGBs to characterise amounts of antibody required to precipitate desired amounts of target protein.

In vitro DNA binding properties of human p53 were examined using an optimised version of a previously developed SMB assay (Palecek et al., 2004), initially to ascertain the dependence on electrostatic forces during binding to various DNA substrates. Human p53-B was incubated with linear dsDNA containing a p53 response element, linear dsDNA containing no response element and ssDNA. It was found that at 50 mM KCl, binding to all DNAs was over 60% and there was no preference towards sequence-specific DNA, activity that is characteristic of p53 (Wolcke et al., 2003). Percentage binding was found to be highest to ssDNA, a feature of p53 that has also previously been documented (Bakalkin et al., 1995; Bakalkin et al., 1994; Selivanova et al., 1996). Increasing the concentration of KCl significantly reduced p53 binding to non-sequence specific dsDNA and 'No DNA' samples, whereas binding to the response element and ssDNA remained consistently high. These results were consistent with the view that p53 can bind to DNA in different modes of varied strength, depending on sequence or structural characteristics. As non-specific protein-DNA interactions are heavily dependent on electrostatic interactions, neutralising charges between

protein and DNA by increasing salt concentrations would decrease protein binding to non-specific DNAs. Whereas p53 binding non-specifically to DNA is determined by the CTD, binding to response elements in DNA is mediated primarily by the CDBD alongside regulation by other domains and is thermodynamically more stable. Recently a study has confirmed that the interaction with consensus DNA sequences enhances p53 conformational stability at physiological salt concentrations, but becomes abolished at 300 mM ionic strength (Ishimaru et al., 2009). Therefore, p53 may also be demonstrating increased binding at higher KCl concentrations to consensus DNA than non-specific dsDNA as the protein may be more stable due to the presence of consensus DNA. As with non-specific p53-dsDNA, binding to non-linear dsDNA, such as ssDNA, is thought to be CTD-dependent. However, the results show that this interaction is considerably stronger than CTD mediated non-specific interactions, indicating that alternative components alongside electrostatic forces are also involved in this interaction.

It was also of interest to detect if p53-B and p53-I exhibited differential DNA binding activity. This was achieved by performing identical assays with both p53s using substrates designed to detect at least 3 modes of p53-DNA interactions – sequence-specific, structure-specific and non-specific. The data obtained from these assays convincingly demonstrated that both p53s had almost identical DNA binding characteristics, suggesting that either both molecules were identical in conformation and modification, or that post-translational modifications that may have occurred in eukaryotic host cells did not significantly influence DNA binding. It was also postulated previously that the apparent difference in electrophoretic mobility of the p53s observed during protein analysis may have arisen if p53s were different isomers. However, if this is indeed the case, it is apparent that both isomers present display similar DNA binding characteristics.

SMB assay results also implied that p53 binding to non-consensus dsDNA was significantly enhanced when an internal hairpin consisting of CTG repeats was present. This interaction was apparently further strengthened as the size of the hairpin structure increased, findings that correlated strongly with a previous study that also detected a strong association of p53 towards identical DNA substrates using an alternative approach (Walter et al., 2005). To confirm this interaction was structure-specific, assays that employed similar substrates that contained alternative triplet repeats were investigated. It was apparent that although p53 bound to a high extent to all hairpin DNAs, there was significant variation that was attributed to sequence. Notably, this sequence-dependence may be apparent due to the effect on the structure of the hairpin, meaning that ultimately this interaction was primarily structure-dependent. This proposal was reinforced by analysis of linear dsDNA containing TRS to

which p53 binding was significantly less, confirming the presence of the hairpin was key to p53 binding. The physiological connotations of such interaction could be substantial as considering the formation of secondary structures in TRS tracts present in certain human hereditary neurological disorders can be multiple-times larger than DNAs investigated in this study, the apparent affinity of p53 towards these structures would be considerable. Such activity may be an example of p53 acting as a global surveillance protein (Albrechtsen et al., 1999; Wiesmuller, 2001) and that maintenance of unstable TRS tracts may be an important activity of p53.

In addition to p53s documented affinity towards DNAs containing hairpin structures, other unorthodox features in DNA have also been reported to enhance p53-DNA binding. Besides ssDNA, insertion-deletion lesion mismatches (Degtyareva et al., 2001; Lee et al., 1995), Holliday junctions (Subramanian and Griffith, 2005), triple-stranded DNA (Dudenhoffer et al., 1998) and even RNA molecules (Riley et al., 2007) are recognised as substrates with which p53 can interact. Thus, the SMB assay was utilised to detect binding to dsDNA containing features including mismatches, insertions, loops and bubble structures. It was found that although binding was generally low to these DNAs relative to previous substrates, a preference towards G/A mismatches, loop and bubble structures was detectable. It was postulated that binding to these substrates was generally lower than anticipated because they were all variants of a dsDNA that was smaller than oligonucleotides previously studied. It was assumed that this decrease in size reduced overall p53 binding, a concept that appeared to be enhanced by comparing non-specific interactions to dsDNAs of varying sizes that also appeared to be DNA length dependent.

For determination of *in vivo* p53-DNA binding, ChIP assays investigating p53 interactions between consensus sequence DNA (p21 and Mdm2 promoter regions) and non-consensus DNA (GAPDH promoter region) were compared in HeLa cells that had, or had not been, subject to UV-irradiation. Data obtained from ChIP assays indicated the presence of an active p53 capable of binding to specific genomic sites in non-UV-irradiated HeLa cells as binding to p21 and Mdm2 regions was significantly higher than GAPDH promoter binding. UV-irradiation did affect p53 binding to all genomic regions although p21 and Mdm2 remained significantly higher than GAPDH promoter binding. Interestingly, precipitation of Mdm2 appeared to increase the most following UV-irradiation, a feature that may represent a change in p53 activity following UV-irradiation that has altered its affinity towards different target sites. However, as Mdm2 is a negative regulator of p53, it is unclear why transactivation of this gene would be beneficial following increased DNA damage that would arise following UV-irradiation.

As significant differences were detected between p53 and Mdm2 consensus promoter regions compared to the GAPDH promoter region that acted as a negative control, the results indicated that ChIP assays were performed successfully. However, another explanation for the observed differences between p53-binding to these regions could be a direct result of the unusual genetics of HeLa cells. Horizontal gene transfer from HPV has dramatically altered the HeLa genome that now demonstrate a hyper-triploid chromosome number, specific numerical deviations and 20 clonally abnormal chromosomes (Macville et al., 1999). These genetic alterations result in the duplication of genes and so, as a consequence, during ChIP assays, certain genes will have a greater probability of being immunoprecipitated as they are present in higher numbers. The influence of this aspect on results obtained during this study remains unknown.

CHAPTER 4

Verification of Prokaryotic Ku and Analysis of DNA Binding

Genomic stability can be greatly compromised by inappropriate DNA double-strand breaks (DSBs) due to their high cytotoxicity. Non-homologous end-joining (NHEJ) is a comparatively error-prone repair pathway that can recognise and ligate together various DSBs independently of homology between DNA duplexes (Hakem, 2008). Whereas eukaryotic NHEJ requires several highly coordinated proteins that each perform specific roles, characterised prokaryotic NHEJ consists of a two-component system comprising Ku proteins and a multifunctional DNA ligase. In contrast to eukaryotic Ku70/Ku80, prokaryotic Ku are believed to function as homodimers that generally only contain domains similar to the central core DNA binding domain and dimerisation region of Ku70/Ku80 (Pitcher et al., 2007). DNA binding analysis of Ku from *Mycobacterium tuberculosis* (MtKu) has shown that prokaryotic Ku exhibit similar activity as eukaryotic Ku since binding affinity is high towards dsDNA ends, but lower towards ssDNA and closed-circular DNA (Weller et al., 2002). *In silico* analysis of the *S. coelicolor* genome has revealed 3 potential homologues of Ku - *SCO5309* (*scku1*), *SCO0601* (*scku2*) and *SCP1.285c* (*scku3*) (Sayer, 2008) that show high similarity to the *M. tuberculosis mtku* gene. Both *scku1* and *scku2* are located on the chromosome of *S. coelicolor*, with *scku1* positioned adjacent to *scprim1* (*SCO5308*) (Sayer, 2008), a gene that shows high homology to the primase region of the *M. tuberculosis LigD* gene. However, *scku3* is located on the plasmid SCP1 and is considerably smaller than *scku1* and *scku2*. Sequence analysis has identified that all 3 *scku* genes show homology with eukaryotic Ku DNA binding domains, but *scku2* also contains a sequence encoding for a domain similar to the SAP domain found on Ku70 that is involved in nucleic acid interactions and chromosomal organisation (Lees-Miller and Meek, 2003).

Prior to investigations into the DNA-binding properties of prokaryotic NHEJ components, recombinant proteins that had previously been cloned into appropriate vectors (Sayer, 2008) were over-expressed and purified by affinity chromatography via N-terminal polyhistidine tags. Proteins were then analysed primarily by SDS-PAGE and western blotting, and ScKu proteins were subjected to analysis by Matrix-assisted laser desorption/ionisation time-of-

flight (MALDI-TOF) peptide mass fingerprint analysis for confirmation of the proteins identities.

Two *in vitro* DNA binding assays were optimised for investigating prokaryotic NHEJ proteins. The SMB assay that had previously been used during human p53 studies was again used, mainly to determine binding affinities towards short double-stranded and single-stranded oligonucleotides. Competition assays were also developed that were initially anticipated to generate data that would confirm preceding SMB results, however, the data obtained initiated new investigations into analysing DNA termini synapsis by ScKu. SMB assays also revealed the high susceptibility of ScKu to proteolysis, a feature speculated to be important in dissociation from repaired DNA. A separate *in vitro* PGB DNA binding assay was also employed to elucidate prokaryotic Ku interactions with various plasmid molecules. Interactions with linear, supercoiled and relaxed-circular plasmid DNA were examined using alternative conditions to verify specific binding preferences of the proteins studied. Importantly, ScKu interactions with oligonucleotides were also investigated using the PGB assay in order to establish comparisons with the previously acquired SMB data.

Experiments were also performed to establish how ScKu proteins influenced activity of a standard DNA ligase not thought to interact with ScKu. Ligation by bacteriophage T4 DNA ligase of linear plasmid DNA with overhang or blunt-ends independent of ScKu were initially analysed which was then compared to ligation with the inclusion of Ku proteins.

4.1 Purification and Analysis of Prokaryotic Ku Proteins.

Full length genes encoding *scku1*, *scku2* and *scku3* that had previously been cloned individually into pET28(a) plasmids (Sayer, 2008) and full length *mtku* that had been cloned into pRB141 (Malyarchuk et al., 2007; Weller et al., 2002) were available for analysis. *In vitro* investigations into the activities of these prokaryotic gene products required the individual over-expression and purification of each recombinant protein. The procedure employed for prokaryotic Ku over-expression was the pET expression system (Section 2.4.1) as high yields of protein can be obtained and this system had previously been used to over-express prokaryotic Ku.

Purification of successfully over-expressed recombinant proteins was performed using affinity chromatography with either Novagen His·Bind® columns or GE Healthcare HisTrap™ HP cartridges. Both techniques utilised the polyhistidine tags that were present on the recombinant proteins. Electrophoretic analysis by SDS-PAGE was used for primary

analysis of the purified protein and western blotting was performed on all recombinant proteins using immunodetection of the polyhistidine tags to confirm the desired protein had been obtained.

4.1.1 Novagen His-Bind® Column Purification.

Initial recombinant Ku purification was achieved using Novagen His-Bind® columns (Section 2.4.3.2) that have a total binding capacity of 10 mg of target protein per column. Cellular extracts from 1 litre cultures of *E. coli* over-expressing the relevant recombinant protein were applied to columns and, following several wash procedures, pure protein was eluted by the addition of imidazole to the columns. Each purification yielded several fractions that contained varying amounts of recombinant proteins and contaminating proteins. All collected fractions were analysed by SDS-PAGE and Coomassie staining (data not shown) to ascertain fractions that contained pure recombinant proteins that were to be subsequently desalted and analysed further.

All recombinant Ku purifications were successful using His-Bind® columns, however yield varied significantly between proteins. The differences in yield may have been a direct consequence of how well each recombinant protein was expressed in *E. coli*. Values for concentrations and amounts of recombinant Ku proteins purified using His-Bind® columns are shown in Table 4.1.

4.1.2 GE Healthcare HisTrap™ HP Purification.

Affinity chromatography was also performed using HisTrap™ HP cartridges in conjunction with an Äktaprime™ plus (Section 2.4.3.3). Although His-Bind® column purification had been successful, to acquire larger quantities of recombinant protein it was decided to employ HisTrap™ HP cartridges as total binding capacity is greater (60 mg protein) and application of protein samples and buffers can be readily controlled via the Äktaprime™. The Äktaprime™ also enabled easy control of elution steps using gradients of increasing imidazole that was predicted to reduce contaminants as only recombinant proteins would elute at a specific percentage of imidazole. It was discovered that using a step gradient with increments of 20 percent provided the most efficient method of elution. Examples of affinity chromatography using HisTrap™ HP and Äktaprime™ plus are shown in Figure 4.1. Yields of recombinant proteins were also improved by increasing volumes of *E. coli* over-expressing

the proteins of interest. Typically, volumes of 3 litres were used for HisTrap™ HP purification.

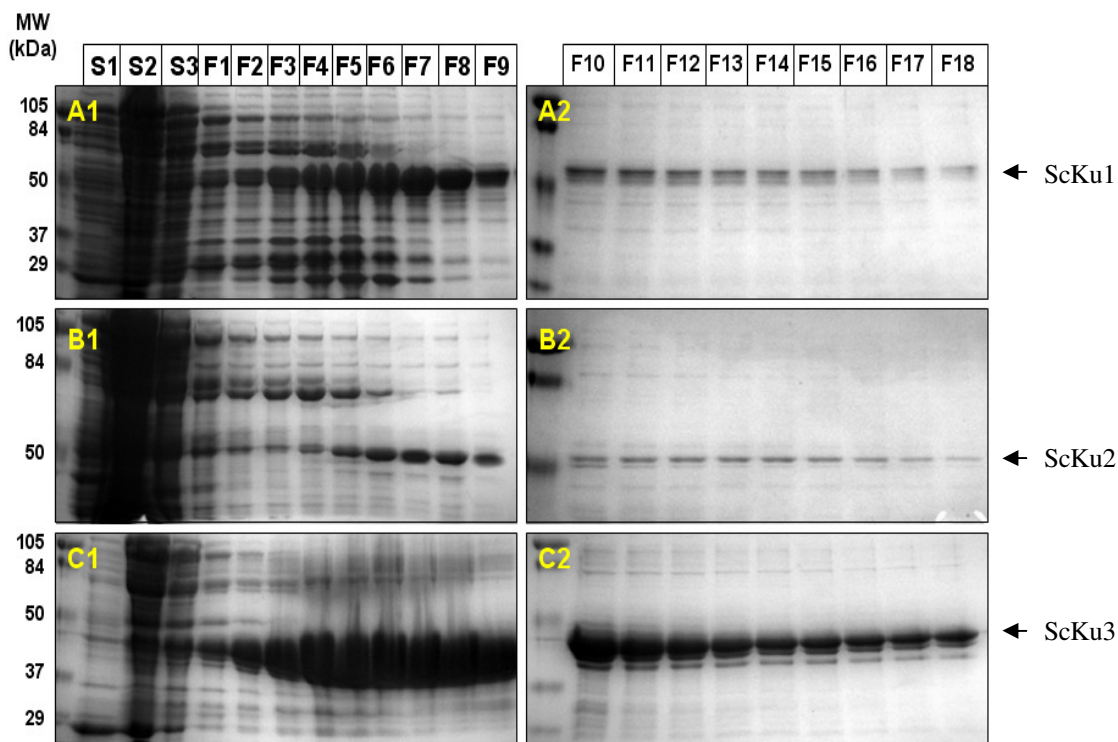


Figure 4.1 SDS-PAGE and Coomassie analysis of fractions collected during HisTrap™ HP purification. 10 µl of fractions were analysed in 12% polyacrylamide gels ran at 200 V for 1 hour. S1 =supernatant eluate, S2 = initial 1 x binding buffer wash, S3 = 1 x wash buffer wash. Fractions relate to eluted pure protein: F1-F18: fractions 1-18. **A1+A2.** Recombinant ScKu1 purification. **B1+B2.** Recombinant ScKu2 purification. **C1+C2.** Recombinant ScKu3 purification.

HisTrap™ HP purification provided greater yields of recombinant protein than previous His·Bind® column chromatography as shown in Table 4.1. This may in part be explained by the increases in *E. coli* cultures used during over-expression, but the higher capacity of the HisTrap™ HP cartridges would also have a considerable influence on recombinant protein yield. Purity of fractions had also been improved by employing a step gradient of increasing imidazole.

Table 4.1 Concentrations and total yields of all recombinant proteins over-expressed and purified from *E. coli*. Values of total protein samples purified via His·Bind® columns are displayed in normal font, and protein samples purified via HisTrap™ HP are in italics.

Protein Name	Concentration (µg/ml)	Concentration (µM)	Yield (µg/L culture)
ScKu1	34	0.9	119
	145	3.7	169
	217	5.6	253
	<i>1434</i>	<i>36.0</i>	<i>1673</i>
	<i>969</i>	<i>25.0</i>	<i>1130</i>
ScKu2	28	0.7	98
	84	2.2	294
	128	3.3	149
	385	9.8	449
	396	<i>10.0</i>	<i>462</i>
ScKu3	160	4.9	560
	215	6.5	753
	455	13.8	531
	<i>4976</i>	<i>150.0</i>	<i>5459</i>
MtKu	192	5.5	671
ScLigD (monomer)	32	0.9	110
ScLigD (dimer)	<i>161</i>	<i>2.3</i>	<i>564</i>

4.1.3 Coomassie and Western Blot Analysis.

Following purification, 2.5 ml of appropriate fractions were desalted and either snap frozen in liquid nitrogen or had glycerol added to a final concentration of 20% and frozen on dry ice. Resulting protein stocks were analysed by SDS-PAGE and Coomassie staining (Figure 4.2).

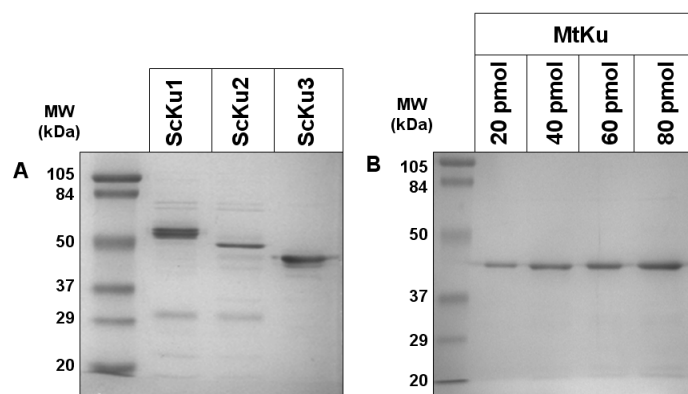


Figure 4.2 SDS-PAGE and Coomassie staining analysis of recombinant prokaryotic Ku proteins. Electrophoresis was performed using 10% polyacrylamide gels analysed at 150 V. **A.** Analysis of 80 pmol of ScKu proteins purified via His·Bind® columns. **B.** Analysis of varying amounts of MtKu proteins purified via His·Bind® columns.

Analysis of the purified proteins (Figure 4.2) showed clear, intense bands representing the purified recombinant protein and low levels of contaminating proteins. Interestingly, electrophoretic migration of all purified Ku proteins was considerably less than expected. ScKu1 (39 kDa), ScKu2 (39 kDa) and ScKu3 (33 kDa) all demonstrated reduced migration, and similarly, MtKu (35 kDa) failed to migrate as predicted. The addition of the polyhistidine tag may reduced electrophoretic migration to an extent, as for ScKu proteins, 6 histidine residues would increase the molecular weight by 0.84 kDa. This, however, does not account for the significant reductions observed. Due to the intensity of the bands present, it was expected that the relevant proteins had been successfully obtained, however, for confirmation, western blots were performed followed by immunodetection specific for the detection of the polyhistidine tags (Figure 4.3).

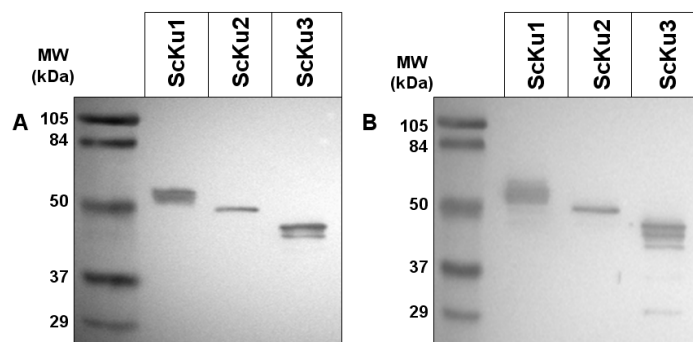


Figure 4.3 Western blot analysis of recombinant ScKu proteins. Proteins were subjected to SDS-PAGE using 10% polyacrylamide gels and western blotting, followed by immunodetection with antibodies specific for polyhistidine tags. A total of 10 pmol of each protein was analysed. **A.** Western blot analysis of recombinant ScKu proteins purified via His·Bind® columns. **B.** Western blot analysis of recombinant ScKu proteins purified via HisTrap™ HP cartridges.

Western blot analysis was consistent with Coomassie results as immunodetection of polyhistidine residues identified bands in identical positions to those previously recognised. This strongly suggested the bands were the required recombinant proteins. It remains unclear why electrophoretic mobility of these proteins was lower than expected. The extra molecular mass of the polyhistidine tag does not fully justify the observed discrepancies, although unusual structural features that may form in this region that are not completely denatured during SDS-PAGE may negatively influence mobility. Similarly, remaining structural features of the Ku proteins would also affect mobility, however increasing SDS concentrations and altering heating procedures did not affect migration.

Alternatively, the unusually high isoelectric points (pIs) of the prokaryotic Ku proteins may also lead to reduced electrophoretic mobility.

4.2 MALDI-TOF Peptide Mass Fingerprinting.

Previous electrophoretic mobility analysis had identified that ScKu proteins failed to migrate according to predicted molecular weights. For further characterisation and to confirm these proteins were as expected, Matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) peptide mass fingerprint analysis was performed on all 3 ScKu proteins.

MALDI-TOF is a powerful procedure that enables the molecular weight determination and structural characterisation of biomolecules. This soft-ionising technique commonly utilises a nitrogen laser beam to analyse biomolecules that are liable to fragment when subjected to conventional ionising methods. For prevention of biomolecule destruction and to facilitate ionisation and vaporisation, a matrix consisting of crystallised molecules such as sinapinic acid are employed that absorb the laser energy and become ionised. The matrix can then assist biomolecule ionisation by transfer of this charge to the molecule whilst simultaneously protecting against the destructive laser energy (Pan et al., 2007). The energy transferred promotes transition of peptide molecules from a solid state into a gaseous state and an electric field accelerates the resulting molecules towards an ion detector where they are detected as an electrical signal. The spectrometer used for detection during ScKu analysis was time-of-flight, which uses a pulsed laser to take individual recordings of ions following MALDI.

Prior to MALDI-TOF, 60 pmol of ScKu1, ScKu2 and ScKu3 were subjected to SDS-PAGE, detected by Coomassie staining and then excised from the gel. Proteins were then treated with trypsin to fragment into smaller peptides that were then analysed by MALDI-TOF. *In silico* peptide mass finger printing (PMF) was employed to compare known protein sequences in various databases with the masses obtained from ScKu peptide analysis. To achieve this, computer software translates the genome of a specific organism into proteins and theoretically digests these proteins into peptides to which the masses of sample peptides are compared (Shevchenko et al., 1996).

Table 4.2 Data obtained from MALDI-TOF PMF analysis for ScKu1, ScKu2 and ScKu3. Experimentally-identified ScKu peptide masses were compared with theoretical masses of *S. coelicolor* protein fragments using the Mascot search engine. Scores above the threshold of 52 are significant ($p < 0.05$) and indicate matched genes. Number of matches of peptides and sequence coverage were also calculated.

Protein	Matched Gene	Mascot Score	Score Threshold	Matched Peptides	Sequence Coverage (%)
ScKu1	SCO5309	164	52	20	60
ScKu2	SCO0601	238	52	27	69
ScKu3	SCP1.285	134	52	20	58

MALDI-TOF spectrums for the 3 ScKu proteins are located in Appendix II. The data obtained from MALDI-TOF PMF (Table 4.2) conclusively established that the proteins analysed were the products of the Ku genes from *S. coelicolor*. The expected genes were found to have Mascot scores considerably higher than the score threshold indicating high similarities. Crucially, no other genes were found to have a score above this threshold. Sequence coverage scores were also very high as, for all ScKu proteins, coverage was between 58 – 69%, further establishing that the ScKu peptides correlated strongly with predicted *S. coelicolor* Ku proteins. As the ScKu proteins were over-expressed in *E. coli*, the *E. coli* genome was also searched to identify the presence of contaminating proteins, however no significant matches were identified.

4.3 *In vitro* SMB DNA Binding Analysis of Prokaryotic NHEJ Components.

Previous investigations into Ku-DNA binding have yielded several interesting, and in some instances conflicting, reports regarding the precise DNA binding activities of this class of proteins. When first identified, the heterodimeric eukaryotic Ku was reported as a dsDNA-end binding protein that had the ability to translocate along DNA (de Vries et al., 1989). Further studies showed that the precise structure of DNA ends was not important and that Ku could interact with hairpin loops and mononucleosomes (Paillard and Strauss, 1991) alongside standard dsDNA termini. The flexibility in Ku-DNA interactions was explored further and it was found Ku binding is largely sequence-independent and is not restricted to linear DNA, but can also bind to nicked-circular DNA and, to a lesser extent, closed-circular DNA and single-stranded circular DNA (Blier et al., 1993). Independent studies confirmed that Ku can directly interact with unusual structures DNA end independently including single-to-double strand transitions and regions of non-complementarity (Falzon et al., 1993). It is generally accepted that eukaryotic Ku has low affinity for ssDNA, however data exists that shows Ku interacts with certain RNA molecules (Kaczmarek and Khan, 1993), implying further the complexity and specificity of eukaryotic Ku. Since the discovery of prokaryotic

Ku homologues, studies have shown high correlation in DNA binding activity between these prokaryotic and eukaryotic counterparts as studies investigating Ku from *M. tuberculosis* found homodimeric Ku bound with high affinity to dsDNA but not to ssDNA or closed-circular DNA (Della et al., 2004; Weller et al., 2002).

Traditionally, *in vitro* Ku-DNA binding has been visualised using techniques including EMSA, protection assays and electron microscopy. A relatively new *in vitro* SMB DNA binding assay (Section 2.7) that had been employed to investigate human p53 DNA binding activity was chosen to investigate the DNA binding activities of the 3 Ku homologues and DNA Ligase D (ScLigD) of *S. coelicolor*. To allow direct-comparison with traditional techniques, the DNA binding activity of the Ku homologue from *M. tuberculosis* was also investigated.

4.3.1 Investigating Ku and DNA Ligase Binding to dsDNA and ssDNA.

As the consensus view regarding Ku-DNA interactions is that Ku binds to dsDNA with significantly greater affinity than ssDNA, it was decided primary investigations would examine the strength with which ScKu proteins bound to dsDNA molecules compared to ssDNA molecules. It is known that DNAs as small as 14 base pairs are sufficient for eukaryotic Ku-DNA binding and so oligonucleotides greater than 30 base pairs (Table 2.7) were expected to be appropriate for studying prokaryotic Ku DNA binding. During SMB assays, 15 pmol of relevant protein was incubated with 20 pmol of available DNA ends, meaning the DNAs would not become saturated and all protein would have available DNA. As it is accepted that multiple Ku can interact with a single DNA molecule, this amount of DNA should be more than adequate to prevent Ku from not having available substrate. Importantly, ionic strength was determined by the concentration of KCl, which was 220 mM during all assays (a value previously used in SMB assays (Section 3.4)). Results from SMB assays investigating ScKu, ScLigD and MtKu interactions with dsDNA and ssDNA are shown in Figure 4.4. Unlike previous SMB experiments using p53, non-specific interactions between prokaryotic NHEJ proteins and SMBs were consistently low (<10%) and therefore are not shown in this section.

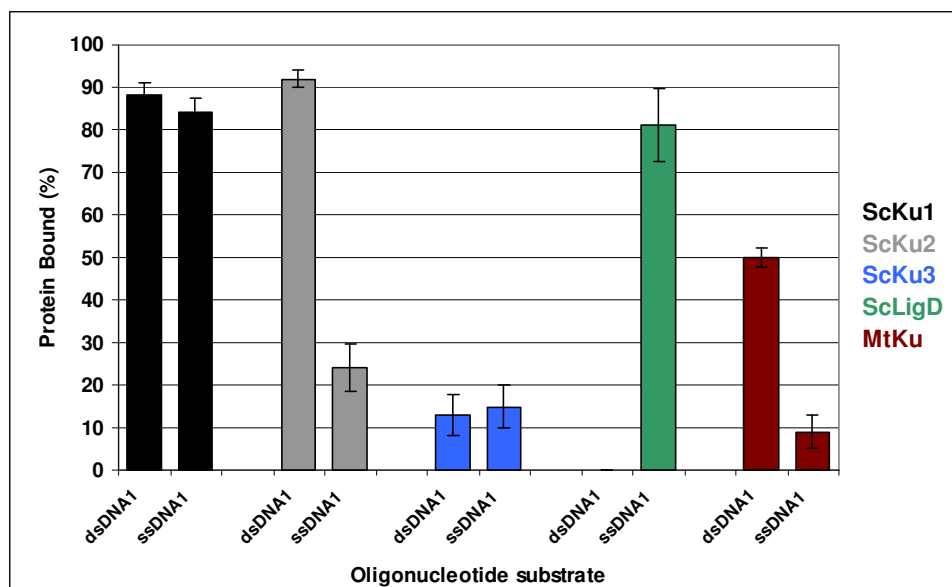


Figure 4.4 *In vitro* SMB DNA binding assay results investigating ScKu, ScLigD and MtKu interactions with dsDNA and ssDNA. Protein samples were collected and analysed as stated (Figure 3.10). Band intensities were measured and values for percentage bound DNA were calculated. Assays were performed at 220 mM KCl. Averages were taken from a minimum of 5 experiments. Standard errors are shown.

The data from SMB assays indicated considerable differences between the prokaryotic Ku proteins being investigated. Considering previously studied eukaryotic and mycobacterial Ku bind preferentially to dsDNA (Dyan and Yoo, 1998; Weller et al., 2002), it was unexpected that ScKu1 demonstrated the ability to bind to ssDNA to a similar extent as dsDNA. Statistical analysis identified there was no significant difference ($t = 0.98$ ($p > 0.05$)) between ScKu1-dsDNA and ssDNA binding. Conversely, ScKu2 demonstrated the predicted activity as binding to dsDNA was significantly greater than ssDNA ($t = 12.6$ ($p < 0.05$)), a result that showed strong similarities with MtKu activity that also bound significantly more to dsDNA ($t = 8.3$ ($p < 0.05$)). Interestingly, ScKu3 bound relatively poorly to both DNA substrates, indicating either this protein has naturally lower affinity towards DNA than other prokaryotic Ku, or the recombinant protein studied was not functioning efficiently. As the results in Figure 4.4 show, ScKu1-DNA and ScKu2-dsDNA binding are very high as almost all protein was precipitated during SMB assays when the appropriate DNA substrate was available. MtKu binding to DNA was generally lower than that observed with ScKu1 and ScKu2, indicating a difference in DNA binding abilities between prokaryotic Ku from different species of bacteria. The binding of ScLigD was also analysed and it was discovered ScLigD could interact with ssDNA, but interacted significantly less to dsDNA ($t = 7$ ($p < 0.05$)), a result that contrasted entirely with ScKu2 and MtKu. Generally, the SMB assay had detected differences in binding activity between the prokaryotic NHEJ components and it was of interest to further characterise affinities of the Ku proteins to dsDNA and ssDNA.

To obtain specific values regarding DNA binding affinity towards dsDNA and ssDNA substrate molecules, SMB assays were performed that investigated binding of prokaryotic Ku to varying concentrations of DNAs. Traditional *in vitro* DNA binding techniques such as EMSA detect binding by detection of the DNA and, therefore, during affinity experiments, the levels of protein are varied as the amount of DNA must remain the same for visualisation. During SMB assays, as it is the protein that is detected, the only component of the assay that can be altered is the DNA. During SMB assays, when 15 pmol of Ku protein were investigated in reaction volumes of 20 μ l the concentrations were 0.75 μ M and 20 pmol of DNAs were at 1.2 μ M (Figure 4.5). To elucidate at which concentrations of DNA prokaryotic Ku binding was at approximately 50%, concentrations of DNA was varied from 0 – 3 μ M and Ku concentrations were maintained at 0.75 μ M. Dissociation constant values (K_d) were calculated using an equation adapted from a previous protein-DNA interaction study (Wieland et al., 2001) detailed in Section 2.7.3.

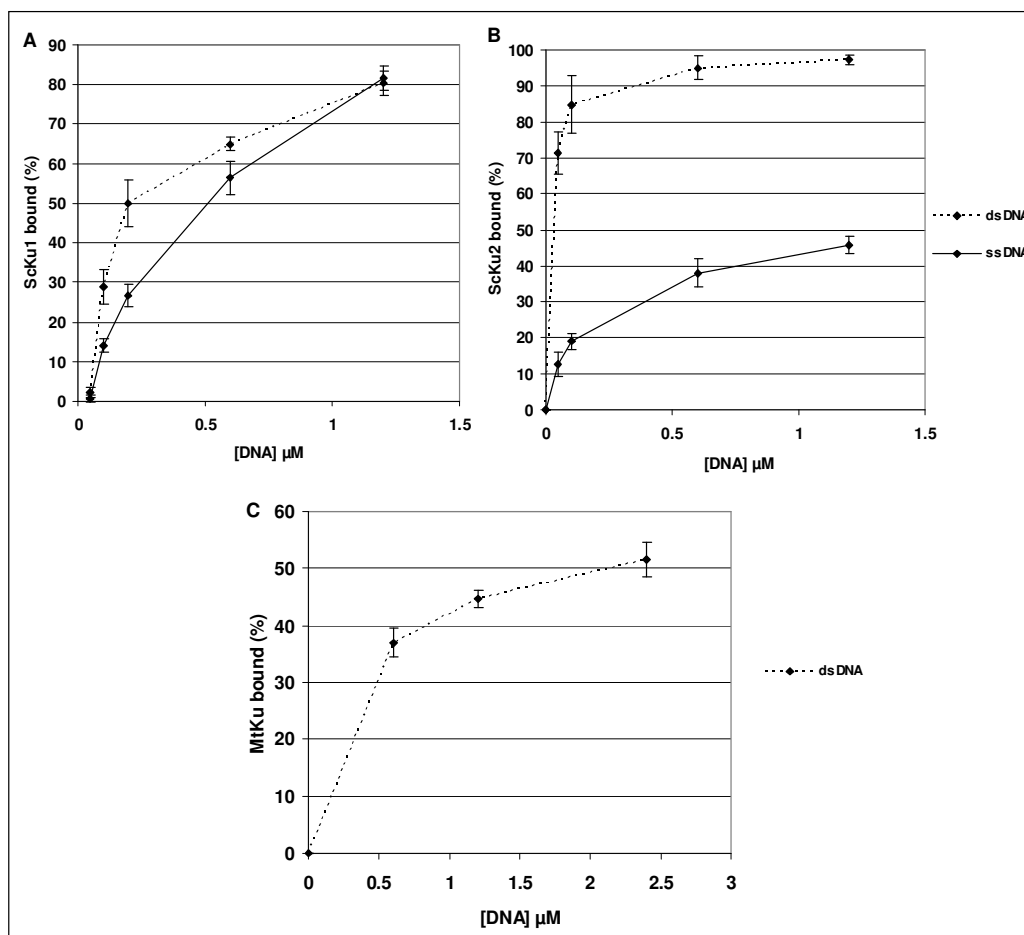


Figure 4.5 *In vitro* SMB assays investigating DNA binding affinities of prokaryotic Ku to dsDNA and ssDNA. Ku concentrations were maintained at 0.75 μ M whereas substrate DNAs were at concentrations between 0 - 3 μ M. Values for K_d were calculated using Equation 6 (Section 2.7.3). Averages were taken from at least 5 experiments. Standard errors are shown. **A.** Analysis of ScKu1 binding to dsDNA and ssDNA. **B.** Analysis of ScKu2 binding to dsDNA and ssDNA. **C.** Analysis of MtKu binding to dsDNA.

Table 4.3 Dissociation constants (μM) for prokaryotic Ku binding to dsDNA and ssDNA. Values were obtained using Equation 6 (Section 2.7.3). NF = No Fit, NDB = No Detectable Binding.

	One Protein / DNA		Two Proteins / DNA		Four Proteins / DNA	
Protein	dsDNA	ssDNA	dsDNA	ssDNA	dsDNA	ssDNA
ScKu1	NF	0.12	0.01	0.30	0.41	0.40
ScKu2	NF	1.63	NF	1.81	NF	1.91
MtKu	1.73	NDB	1.91	NDB	2.01	NDB

Note: “No Fit” refers to data where the DNA binds to the protein too tightly to enable a K_d to be obtained using the model of 1, 2 or 4 protein molecules bound per DNA.

As the precise mode of the prokaryotic Ku-DNA binding remains uncertain, dissociation constant values were calculated for 1, 2 and 4 protein binding per DNA molecule (Table 4.3). It was soon recognised that the values obtained for particular proteins when analysed as monomeric proteins were not practical as negative results were produced as a consequence of very tight protein-DNA interactions. This was true for ScKu1 and ScKu2 when dsDNA was used as a substrate, however when dimeric or tetrameric protein binding is assumed, values can be obtained for ScKu1. This occurrence may imply that multiple proteins were able to bind to single DNA molecules used during these assays and, therefore, the single-protein binding model does not fit for the ScKu proteins investigated. As MtKu has previously been shown to exist as a homodimer in solution (Weller et al., 2002), it is likely the value obtained from calculations assuming dimeric binding is the most accurate value.

4.3.2 ScKu Competition Assays.

Competition assays can be used to distinguish a protein’s preference between DNA substrates. Such assays are especially useful when affinity towards different DNAs has been found to be similar or to identify if a protein already bound to DNA can be removed and interact with a separate DNA molecule. Previous SMB assays had detected that ScKu1 and ScKu2 bound with varied affinity towards dsDNA and ssDNA molecules, and it was anticipated competition assays investigating the same DNAs could provide more evidence to support the SMB data. Assays were originally performed whereby ScKu1 and ScKu2 were initially incubated with dsDNA and then following removal of unbound Ku, either competing unbiotinylated dsDNA or ssDNA was added to the reactions. Unbiotinylated DNAs were then removed in the subsequent wash step alongside any protein now bound to the free DNAs, and as with previous SMB assays, protein bound to biotinylated DNA were also removed and analysed. Therefore, for each reaction, 3 samples were collected: unbound protein from the primary incubation with DNA-SMBs, protein bound to competing DNA and

protein bound to DNA-SMBs. From the resulting analysis, percentages of protein removed following competition were calculated. Surprisingly, irrespective of Ku protein or competing DNA, only very low amounts of ScKu protein bound to competing DNA was detectable, even when competing DNA was 3 times in excess of the DNA-SMBs (data not shown). It was feasible that once bound to DNA, ScKu proteins may have been difficult to displace and so the addition of competing DNA following initial DNA incubations was ineffective.

To further assess the effect of competing DNA, assays were performed whereby competing DNA was added concurrently with protein to the reactions. This provided DNA-SMBs and free unbiotinylated DNAs with equal opportunities for ScKu binding and any interactions with competing DNA would not require ScKu dissociation from DNA.

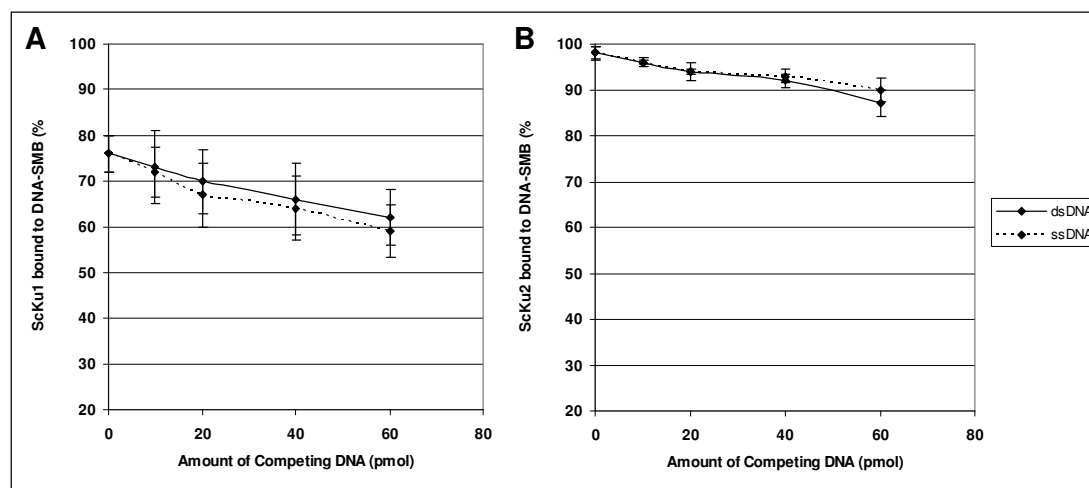


Figure 4.6 *In vitro* SMB DNA binding competition assays investigating ScKu1 and ScKu2 interactions to dsDNA and ssDNA. Competing DNA was added simultaneously with protein to ensure no bias towards DNA attached to SMBs. The amount of DNA-SMBs was maintained at 20 pmol and amounts of competing DNA varied from 0 – 60 pmol. Reactions were performed at 220 mM KCl. Averages were taken from 3 experiments. Standard errors are shown. **A.** ScKu1 competition between dsDNA-SMBs and dsDNA or ssDNA. **B.** ScKu2 competition between dsDNA-SMBs and dsDNA or ssDNA.

It would be expected that if a DNA-binding protein was incubated with equal amounts of 2 DNAs of the same structure and sequence that binding to both DNAs would be approximately 50%. As shown in Figure 4.6, the effect of the addition of competing DNA to reactions investigating ScKu1 and ScKu2 binding to DNA was considerably lower than anticipated. During ScKu1 assays, when either dsDNA or ssDNA competing DNA was added in equal amounts as dsDNA-SMBs, binding to dsDNA-SMBs was reduced by less than 15% and when competing DNA was 3 times in excess of DNA-SMBs, binding to DNA-SMBs was reduced by less than 20%. Correspondingly, the effect of competing DNA on ScKu2 binding was

even lower as when competing DNA was 3 times in excess, binding to DNA-SMBs was decreased by less than 15%.

To ascertain the apparent preference that ScKu had for DNA-SMBs over free DNAs, several investigations were performed that would identify why the ScKu proteins were discriminating between DNAs. An important factor to ascertain was if unbiotinylated DNA could become physically associated with SMBs and therefore assays were repeated without the inclusion of biotinylated DNA, meaning any protein precipitated with SMBs would be dependent on free DNA physically associated with SMBs. The results from these experiments revealed no ScKu were precipitated (data not shown) and, therefore, free DNA was not attaching to SMBs.

Alternatively, it was proposed that the competing DNA may have been associating with DNA-SMBs either via direct interactions with the biotinylated DNA, or through bound ScKu proteins that theoretically may juxtapose DNA termini of free DNA and DNA-SMBs. This aligning of DNA termini may have been an example of ScKu replicating an essential function performed during NHEJ whereby a synapse is formed between 2 DNA termini of a DSB to form a stable intermediate for intermolecular ligation. If true, free DNA may have been precipitated during SMB assays and this would explain the results obtained during competition assays. In a similar technique to a previous study that successfully detected DNA end aligning by Ku proteins (Ramsden and Gellert, 1998), ScKu proteins were incubated with biotinylated DNA and excess amounts of unbiotinylated ssDNAs modified with 5' fluorophores were added to reactions. Theoretically, if ScKu were aligning DNA-SMB with free DNA, free DNA would remain associated with SMB-DNA-ScKu complexes and this would be indicated by an increase in fluorescence emission in these reactions following wash steps. To confirm this effect was ScKu dependent and not due to interactions involving only DNAs, simultaneous reactions without the inclusion of ScKu proteins were also performed.

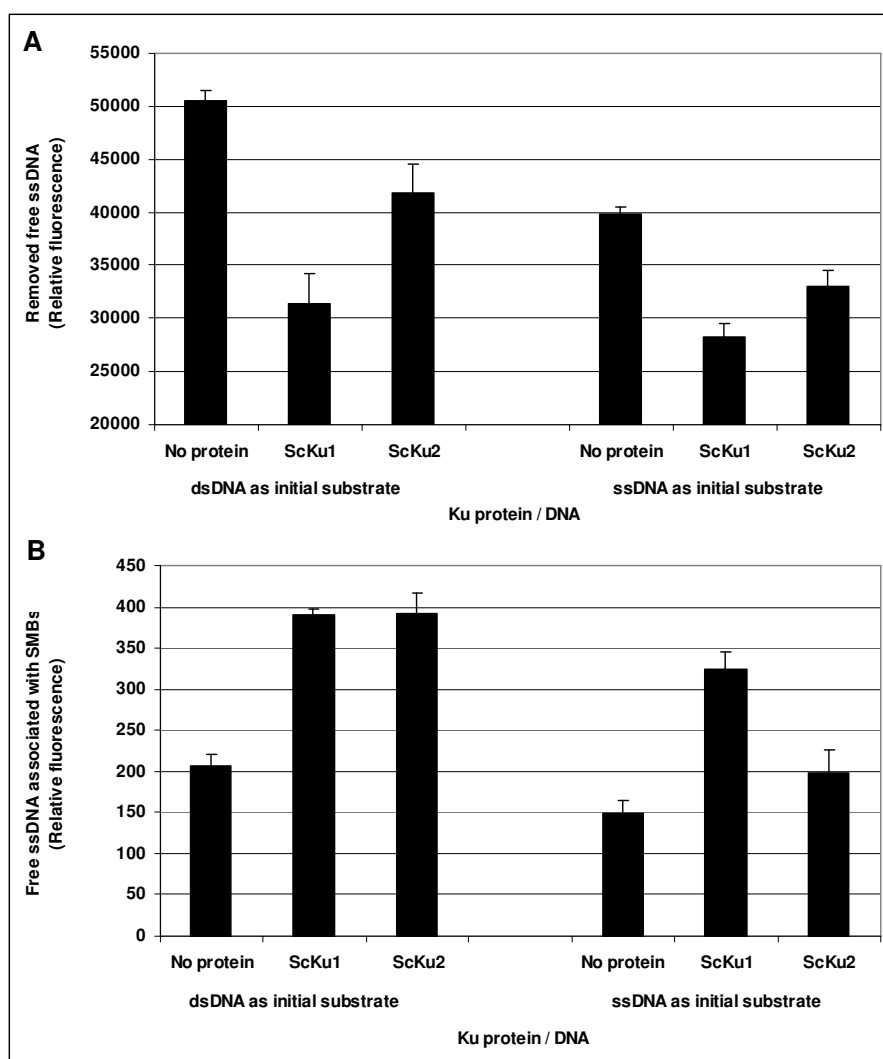


Figure 4.7 Determining DNA termini alignment by ScKu1 and ScKu2 using modified SMB DNA binding assays. ScKu proteins incubation with DNA-SMB was followed by addition of unbiotinylated fluorescently labelled ssDNA. Following wash steps, any fluorescence detected may be caused by ScKu proteins forming synapsis between biotinylated and unbiotinylated DNAs. Averages were taken from 6 experiments. Standard errors are shown. **A.** Fluorescence emission from ssDNA removed after incubation with DNA-ScKu. **B.** Fluorescence emission from ssDNA precipitated with SMB-DNA-ScKu complexes.

The data collected from ScKu DNA termini alignment investigations indicate that the presence of both ScKu1 and ScKu2 did affect the amount of free ssDNA precipitated with SMBs (Figure 4.7 A). This difference is most apparent when dsDNA is used for initial ScKu binding as significantly less fluorescent ssDNA is removed from reactions when ScKu1 is present ($t = 6.4$ ($p < 0.05$)) and also significantly less fluorescent ssDNA is removed when ScKu2 is present ($t = 5.0$ ($p < 0.05$)). The implication that less fluorescent ssDNA is removed when ScKu are bound to dsDNA is enhanced in Figure 4.7 B as significantly more fluorescent ssDNA has been precipitated when ScKu1 or ScKu2 is attached to dsDNA ($t = 12.5$ ($p < 0.05$) and $t = 6.6$ ($p < 0.05$), respectively). Results investigating free ssDNA

precipitation following ssDNA-SMB incubation are similar to dsDNA-SMB data, however ScKu2 appears to be less efficient in aligning ssDNA molecules. Although Figure 4.7 A shows that there is no significant difference between removed ssDNA fluorescence emission between ScKu1 or ScKu2 ($t = 1.5$ ($p > 0.05$)) when initially bound to ssDNA-SMB, data in Figure 4.7 B illustrates that significantly less ssDNA is precipitated when ScKu2 is bound to ssDNA-SMB than ScKu1 ($t = 8.7$ ($p < 0.05$)). These findings indicate ScKu1 is more efficient at aligning ssDNA termini compared to ScKu2, which correlates with previous SMB data that show ScKu1 binds to ssDNA with higher affinity. However, ScKu2 does appear to increase precipitation of ssDNA even when initially bound to ssDNA-SMB, suggesting that although this protein may have relatively low affinity for ssDNA, interactions can still be detected.

Results from SMB assays investigating synopsis of DNA termini by ScKu proteins may explain, to an extent, why competing DNAs in previous assays decreased binding to DNA-SMB by relatively low amounts. However, the amounts of fluorescent DNA precipitated when ScKu were present were not as considerable as may have been expected. This is reinforced by the substantial differences in fluorescence between removed free ssDNA and precipitated free ssDNA. For example, only 1.3% of free ssDNA is precipitated by ScKu1 when initially bound to dsDNA-SMB. One explanation for the relatively low amounts of precipitated DNA during synopsis SMB assays compared to results from competition assays is that the fluorophore present on free ssDNA molecules may have had a considerable effect on ScKu function. Whereas free DNAs in competition assays have 2 free-ends available for alignment, fluorescent DNA only has 1 available end. Therefore, only a single fluorescent DNA can be aligned with DNA-SMB. In contrast, during competition assays several free DNAs can be aligned together to create complexes consisting of several DNA molecules with the capacity for precipitation (Figure 4.17).

4.4 *In vitro* Protein-G Bead (PGB) Assay.

The *in vitro* SMB DNA binding assay had revealed differential activity between the 3 ScKu proteins of *S. coelicolor*, indicating that these proteins may have alternate cellular functions, potentially independent of NHEJ. To further characterise prokaryotic Ku-DNA interactions, an additional *in vitro* DNA binding assay was employed that captures DNA via immunoprecipitated proteins attached to Protein-G beads (PGB) (Section 2.8). In contrast to the SMB assay, this assay is not restricted to biotinylated oligonucleotides, allowing larger DNA molecules to be used as binding substrates. This approach considerably expands the scope of assessment of protein-DNA interactions as molecules exhibiting diverse

conformations such as linear, supercoiled, nicked-circular and relaxed-circular DNA can be investigated. As DNA topology is known to significantly influence protein-DNA interactions, this assay may provide additional information regarding prokaryotic Ku DNA binding activity. The *in vitro* PGB assay used in this study is an optimised version of an assay previously used to investigate p53 binding to various target plasmid DNAs. In these previous studies, plasmid DNA was added in excess of protein and only captured DNA was analysed, meaning calculations of precise values for percentage DNA captured was not possible. However, during this study, protein was added in excess of DNA so that saturation of protein was prevented and all DNA could be precipitated. Consequently, direct comparison of both un-captured and captured plasmid DNA could be used to calculate percentages of DNA precipitated. As described, KCl concentrations were also altered appropriately to reflect more realistic cellular conditions.

4.4.1 Exploring the Rapid Degradation Properties of ScKu Proteins.

Prior to PGB DNA binding assays, it was discovered that in the presence of plasmid DNA, ScKu proteins appeared to vanish following SDS-PAGE and western blot analysis. Subsequent analysis by Coomassie staining identified that this disappearance was a direct result of rapid degradation of ScKu proteins by an unknown component of the plasmid preparation (Figure 4.8).

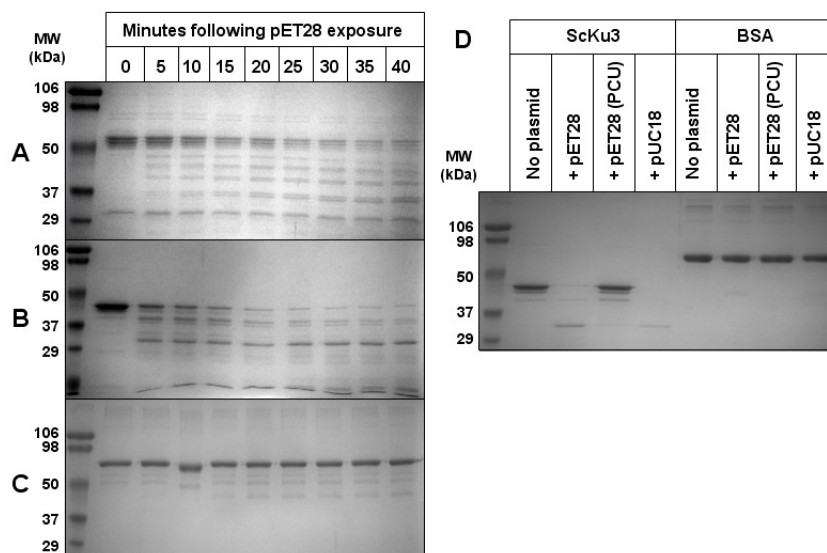


Figure 4.8 SDS-PAGE and Coomassie staining analysis of ScKu and BSA protein degradation following pET28 and pUC18 exposure. Plasmid sample (1 μ l of 100 ng / μ l) was added to 60 pmol of ScKu or BSA in a final volume of 30 μ l and incubated for up to 40 minutes. Proteins were analysed by SDS-PAGE using 10% polyacrylamide gels. **A.** ScKu1 degradation following pET28 exposure. **B.** ScKu3 degradation following pET28 exposure. **C.** BSA degradation following pET28 exposure. **D.** ScKu3 and BSA degradation following pET28, pET28 post clean-up (PCU) or pUC18 incubation for 40 minutes.

As Figure 4.8 shows, both ScKu1 and ScKu3 were degraded considerably faster than BSA when exposed to plasmid, suggesting a component of the plasmid sample was responsible for the disappearance of ScKu during plasmid competition assays. It can be seen that after 5 minutes, both ScKu1 and ScKu3 show visible signs of degradation and the formation of smaller products as regular bands of lower molecular weight. Although ScKu2 demonstrated higher degradation than BSA, it was found to be more stable than ScKu1 and ScKu3 following plasmid exposure (data not shown). The degradation may have resulted from protease activity to which the ScKu proteins appeared highly susceptible. To ascertain if ScKu degradation could be prevented, plasmid preparations were purified further (Section 2.2.4) and added to ScKu samples to identify if the source of degradation had been removed (Figure 4.8 D). It was found purification of plasmid preparations was sufficient to prevent ScKu degradation and that protease activity was likely to have caused the observed fragmentation. In separate experiments, degradation of a polyhistidine tagged recombinant MtrA protein was also investigated, however no protein reduction was observed (Appendix III), further indicating ScKu are unusually prone to degradation. Potentially, the rapid fragmentation of ScKu proteins detected may be crucial in the removal of ScKu from DNA following NHEJ. Whereas most components of the NHEJ apparatus can disassemble from ligated DNA and leave via passive diffusion, the Ku dimer is thought to be left encircling the DNA and might require proteolytic cleavage for removal (Schar et al., 1997; Walker et al., 2001). If ScKu are indeed highly susceptible to protease activity as implied from these results, the removal from DNA by proteolytic cleavage may now have additional foundation.

4.4.2 ScKu-Plasmid and Oligonucleotide DNA Interactions.

Preliminary assays focused on ScKu interactions with linear and supercoiled pUC18 at varying KCl concentrations. Results from SMB assays reiterated the important role that ionic strength plays during protein-DNA interactions and so it was important to establish how KCl concentration affected interactions in the PGB assay. The plasmid pUC18 was digested with appropriate restriction enzymes to create linear DNA with blunt-ends or 5' linear overhang-ends, or left undigested as a supercoiled molecule. These 3 species of pUC18 were then used as ScKu1, ScKu2 and ScKu3 DNA binding substrates during PGB assays. Plasmid DNA that was captured by protein and subsequently immunoprecipitated was subject to electrophoretic analysis alongside DNA that was not immunoprecipitated (Figure 4.9), allowing for determination of the percentage of DNA bound to the Ku protein present. To obtain precise values for the amounts of DNA captured or not captured, the ImageJ software was used to measure band intensities (Figure 4.10). Typically, 400 ng of Ku protein and 250 ng of plasmid DNA were analysed during each assay.

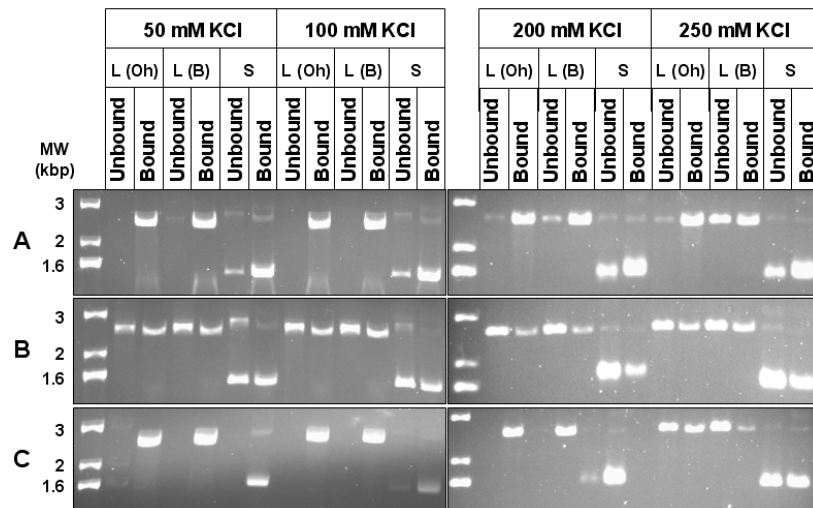


Figure 4.9 Electrophoretic analysis of unbound and bound plasmid DNA following an *in vitro* PGB assay investigating ScKu-DNA interactions. In each experiment 400 ng of relevant ScKu protein was used to capture 250 ng of either linear pUC18 with overhangs (L (Oh)), linear pUC18 with blunt ends (L (B)) or supercoiled pUC18 (S) with varying [KCl]. Bound and unbound DNA was identified after electrophoresis on 0.8% agarose. **A.** Plasmid DNA from ScKu1 PGB assays. **B.** Plasmid DNA from ScKu2 PGB assays. **C.** Plasmid DNA from ScKu3 PGB assays.

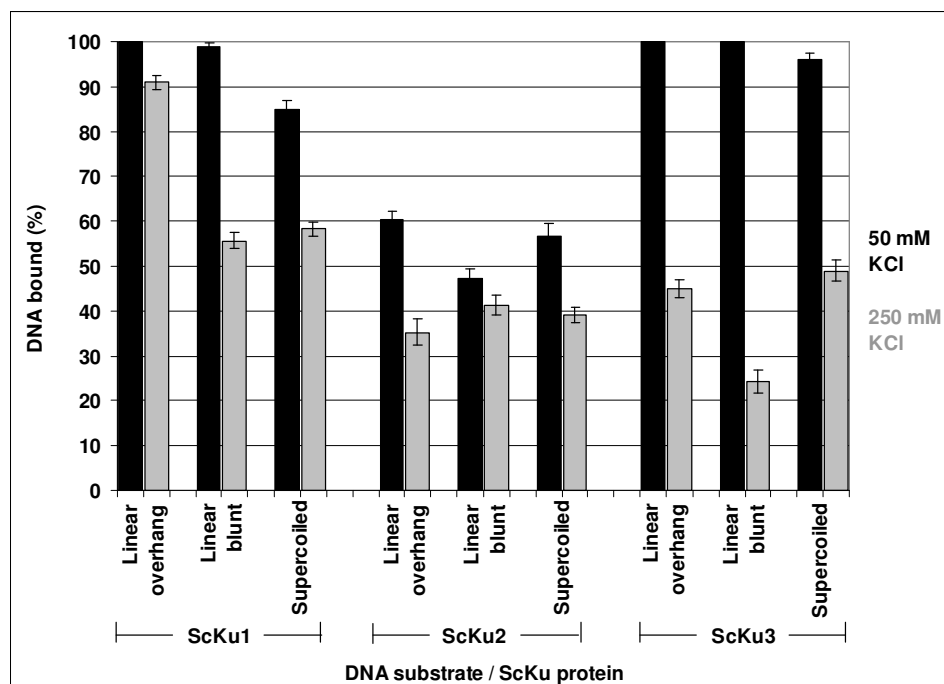


Figure 4.10 Calculated values of plasmid DNA collected from PGB assays described in **Figure 4.9**. Band intensities following electrophoretic analysis were measured using ImageJ software. Averages were taken from 3 experiments. Standard errors are shown.

The results obtained from ScKu-plasmid interactions again identified variation in activity between the 3 Ku proteins. Generally, binding to plasmid DNA was very high, even when the concentration of KCl was increased to 250 mM (a concentration of monovalent cations

greater than natural cellular conditions). ScKu1 and ScKu3 had higher DNA binding activity than ScKu2, which was unexpected as ScKu3 had previously shown relatively low DNA binding activity when oligonucleotides were used as substrates, and ScKu2 had been previously shown to bind with high affinity towards dsDNA. It is possible that ScKu3 may show significantly greater DNA binding activity towards plasmid DNA over oligonucleotides and that it may only interact with larger DNA molecules. As previous studies found Ku only required 14-18 base pairs for DNA binding to occur (Falzon et al., 1993), the differences observed may be dependent upon alternative factors such as experimental conditions.

As expected, it can be seen that increases in KCl concentration directly influenced DNA binding of all 3 ScKu. As non-specific protein-DNA interactions are less favourable at higher ionic strength, only specific interactions remain. This effect is apparent for ScKu binding to all plasmid DNAs as, when KCl is increased, binding becomes reduced; however, for specific interactions this reduction is significantly less. Consequently, it is apparent that both ScKu1 and ScKu3 show a preference towards linear DNA with overhangs compared to linear DNA with blunt-ends, as at 250 mM KCl, binding to blunt-ends is reduced compared to overhang-end binding (ScKu1 - $t = 10.2$ ($p < 0.05$), ScKu3 - $t = 3.1$ ($p < 0.05$)). This preference is not apparent with ScKu2, indicating the nature of the DNA termini do not influence ScKu2 DNA interactions.

Interestingly, all 3 ScKu proteins demonstrated the ability to interact with supercoiled plasmid, a characteristic previously not associated with eukaryotic or prokaryotic Ku (Dynan and Yoo, 1998; Weller et al., 2002). At 50 mM KCl, ScKu1 binds to supercoiled DNA significantly less than linear DNA with blunt-ends ($t = 4.7$ ($p < 0.05$)) but at 250 mM KCl this difference is no longer apparent. However, binding to overhang linear DNA remains significantly higher than supercoiled binding at 250 mM KCl ($t = 12.4$ ($p < 0.05$)). These results imply ScKu1 binds most efficiently to linear DNA (with a preference towards linear DNA with overhang-ends), but can also interact to a lesser extent with supercoiled DNA. ScKu3 demonstrates alternate characteristics as there is no significant difference between binding to linear or supercoiled DNA at 50 mM KCl ($t = 1.9$ ($p > 0.05$) (comparing blunt-end linear and supercoiled DNA binding)) and, surprisingly, ScKu3 binds preferentially to supercoiled DNA over blunt-end DNA at 250 mM KCl ($t = 3.6$ ($p < 0.05$)), indicating that ScKu3 preferentially binds to linear DNA with overhang-ends and least favourably to linear DNA with blunt-ends. In contrast, ScKu2 does not appear to have any considerable preference towards the different plasmid molecules though measurable differences are evident.

It is generally accepted that initial binding of Ku molecules to DNA occurs via free DNA termini and so interactions with closed-circular molecules is thought to be minimal. The results in Figure 4.10 indicate all ScKu proteins are able to interact reasonably tightly with supercoiled DNA as, even at 250 mM KCl, in some instances binding was as high as approximately 60%. These results were unexpected as supercoiled DNA would not have available free-ends to thread through the Ku dimer molecule and so ScKu may have an alternative mode of interacting with DNA.

The *in vitro* PGB DNA binding assay had been used to successfully detect ScKu-DNA interactions and identified that the 3 ScKu proteins had varied DNA binding activity. The *in vitro* SMB DNA binding assay had also detected varied activity between these proteins and so it was of interest to see if both assays would reveal comparative results if the same DNA molecules were used as binding substrates. It was not possible to use plasmid DNA during the SMB assay as no biotinylated plasmid was available. However, due to the flexibility of the PGB assay, oligonucleotides could be used instead of plasmid molecules. To this effect, assays were performed using identical amounts of ScKu and oligonucleotide DNA as described in SMB assays. Samples of bound and unbound oligonucleotide DNA following individual assays were analysed by native electrophoresis, membrane transfer and subsequent chemiluminescent detection of biotinylated tags. It was discovered during initial assays that Ku binding to oligonucleotides was not as efficient as plasmid DNA, since analysis of bound and unbound oligonucleotides revealed significantly more unbound DNA in relation to bound DNA. Consequently, it was not possible to analyse both bound and unbound DNA simultaneously due to the differences in intensity following chemiluminescent detection. Therefore, only oligonucleotides precipitated during PGB assays were subjected to analysis.

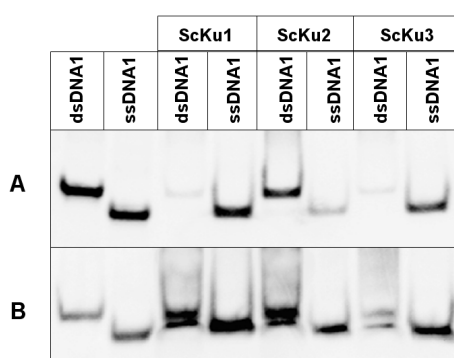


Figure 4.11 Electrophoretic analysis and chemiluminescent detection of precipitated oligonucleotide DNA following *in vitro* PGB assays investigating ScKu-DNA interactions. Electrophoresis was performed at 50 V using 12% native polyacrylamide gels. Markers relating to dsDNA and ssDNA are shown. **A.** Initial assay using identical oligonucleotide substrates used during SMB assays (dsDNA produced with excess unbiotinylated ssDNA). **B.** Example of a PGB assay investigating ScKu binding to dsDNA (produced with equal amounts of biotinylated and unbiotinylated ssDNA) and ssDNA.

As shown in Figure 4.12 A, it was apparent that there were considerable inconsistencies between both *in vitro* assays. Results from SMB assays identified that ScKu1 bound to dsDNA and ssDNA with similar affinity, ScKu2 had significantly higher affinity towards dsDNA over ssDNA and ScKu3 demonstrated relatively low DNA binding activity to any oligonucleotide (Figure 4.4). In contrast, only the ScKu2 data from PGB assays was consistent with SMB assay data. The apparent findings that ScKu1 bound weakly to dsDNA and that ScKu3 could interact with ssDNA contradicted the SMB findings. It was realised that a significant difference between the 2 assays was that unbiotinylated oligonucleotides that were removed during SMB experiments would still be present in PGB assays. During production of dsDNA oligonucleotide substrates (Section 2.3.1), to ensure all biotinylated ssDNA were incorporated into dsDNA molecules, unbiotinylated ssDNA were added in excess. These are removed during SMB wash steps but would remain present during PGB assays. Therefore, it was anticipated that ScKu1 binding to dsDNA was relatively weak compared to ssDNA in PGB assays because excess amounts of unbiotinylated ssDNA were also present in the reaction and these would compete for ScKu1 molecules. It was also thought that ScKu2 was not affected by the excess ssDNA because ScKu2-ssDNA binding had previously been detected as relatively low. To test this, dsDNA were produced at a ratio of 1:1 (biotinylated: unbiotinylated) in order to reduce ssDNA molecules and assays were repeated. As shown in Figure 4.11 B, considerably more dsDNA was precipitated by ScKu1 than previously detected suggesting that, as predicted, the excess contaminating ssDNA was directly competing for ScKu1. ScKu2 DNA binding activity remained constant as relatively more dsDNA was precipitated compared to ssDNA. Again ScKu3 appeared to bind to ssDNA, which was previously not detected during the SMB assay. Apart from the apparent ability of ScKu3 to interact with ssDNA, the results from PGB assays investigating ScKu binding to oligonucleotide DNA demonstrated high agreement with SMB assay results.

4.4.3 The Temperature Sensitive DNA Binding Activity of MtKu.

The unexpected finding that ScKu proteins bound efficiently to closed supercoiled DNA lead to speculation about whether MtKu would bind to similar DNAs. MtKu DNA binding activity has previously been characterised and it was identified that it bound tightly to dsDNA but not to ssDNA or closed-circular DNA (Weller et al., 2002). Initial PGB assays were performed to investigate binding to linear pUC18 containing blunt or overhang ends under identical conditions to ScKu assays. The results indicated that MtKu binding was considerably lower than ScKu as less than 10% of linear plasmid DNA was precipitated. As MtKu had previously demonstrated DNA binding activity during *in vitro* SMB assays, it was expected MtKu would interact with linear pUC18 and so variations in assay conditions were

thought to account for the lack of activity. One considerable difference between the assays was temperature: whereas SMB assays were performed at 30 °C, due to the longer incubation periods and to help maintain assay component stability, PGB assays were performed at 4 °C. To determine the temperature-dependence of MtKu, PGB assays were performed at temperatures between 4 – 40 °C (Figure 4.12 A).

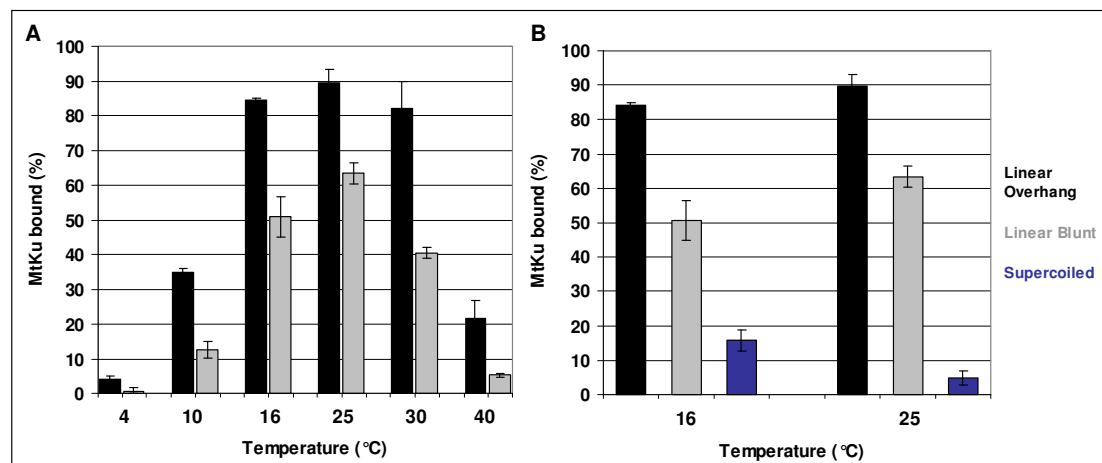


Figure 4.12 Using the PGB DNA binding assay to determine MtKu-DNA interactions. Data was collected from 3 *in vitro* PGB assays that were investigating MtKu binding to linear and supercoiled pUC18 DNAs at varying temperatures and at 50 mM KCl. Band intensities regarding non-precipitated and precipitated DNA following electrophoretic analysis were measured using the computer software ImageJ and percentage bound values were calculated. Standard errors are shown. **A.** MtKu binding to linear pUC18 containing blunt or overhang ends at various temperatures. **B.** MtKu binding to linear and supercoiled pUC18 at 16 °C and 25 °C.

The results in Figure 4.12 A show that MtKu DNA binding activity is heavily dependent upon temperature and that DNA-binding was most efficient at approximately 25 °C. At temperatures less than 16 °C and more than 40 °C binding was considerably lower, indicating that, MtKu has a narrower temperature range at which it is active than apparent with ScKu proteins. Temperature can affect protein structure and flexibility by inducing changes in bond energies, and proteins are generally more flexible at higher temperatures. This increased flexibility may be crucial in facilitating contact between MtKu and DNA and that this is not possible at lower temperatures where the protein may be more rigid and less conformationally favourable.

It was also of interest to determine if MtKu could bind to closed DNA molecules, and so investigations were performed at optimum temperatures to assess this activity. As Figure 4.12 B shows, binding to supercoiled pUC18 was significantly less than binding to linear DNA with the lowest percentage binding (blunt-end linear) (16 °C – $t = 5.3$ ($p < 0.05$), 25 °C

– $t = 17.6$ ($p < 0.05$)). In contrast to ScKu, MtKu binding to supercoiled pUC18 was very low, and binding to plasmid DNA in general was also lower as reductions in ionic strength were required for interactions to occur. These results are consistent with previous MtKu studies (Weller et al., 2002) and, therefore, suggest that considerable differences exist regarding DNA binding activities between MtKu and ScKu proteins.

4.4.4 Prokaryotic Ku Interactions with Relaxed Circular DNA.

Unlike MtKu, ScKu proteins had been found to interact efficiently with closed supercoiled DNA molecules, however it was not clear which features of the DNA allowed this interaction to occur. As supercoiled DNA frequently contains regions of secondary structures, including hairpins, it was postulated these unusual features may govern the observed interactions. Therefore, it was of interest to discover if ScKu could still interact with circular DNA molecules that contained far fewer secondary structures. The restriction enzyme BbvC1 cleaves the asymmetric DNA sequence 5'-CC↓TCAGG-3' / 5'-GC↓TGAGG-3' as indicated by the arrows. Unlike many Type II restriction enzymes, BbvC1 consists of 2 different subunits that contain catalytic sites for DNA strand hydrolysis that act independently and strand-specifically. BbvC1 mutants deficient in 1 subunit exist that cleave only 1 strand of the recognition sequence (Bellamy et al., 2005; Heiter et al., 2005) and, therefore, can nick substrate supercoiled dsDNA to produce relaxed-circular molecules. Plasmid pL1, a derivative of pUC19 (Gowers et al., 2005), was available that contained the BbvC1 recognition site. This plasmid was subjected to mutant BbvC1 (New England BioLabs) digestion to produce the desired relaxed-circular DNAs that were then used as DNA binding substrates for prokaryotic Ku during PGB assays (Figure 4.13).

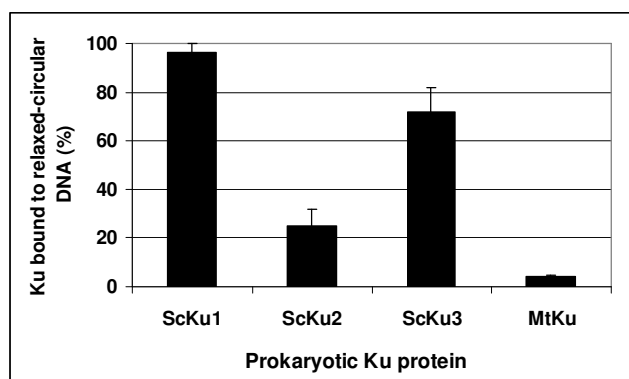


Figure 4.13 Calculated percentages of relaxed-circular plasmid DNA precipitated by prokaryotic Ku proteins during *in vitro* PGB assays. ScKu assays were performed at 200 mM KCl, MtKu assays were performed at 50 mM KCl. Band intensities regarding non-precipitated and precipitated DNA following electrophoretic analysis were measured using the computer software ImageJ and percentage bound values were calculated. Averages were taken from at least 3 experiments. Standard errors are shown.

Results from investigations into prokaryotic Ku binding to relaxed-circular DNA demonstrated high similarities with Ku binding to supercoiled DNA as both ScKu1 and ScKu3 bound tightly to the DNA, whereas ScKu2 and MtKu bound relatively poorly. As both ScKu1 and ScKu3 could interact with circular DNA lacking secondary features, the binding observed during supercoiled substrate assays may not have been dependent upon unusual structures in the DNA. As mentioned, Ku has previously been shown to have high affinity towards circular DNA containing nicks (Blier et al., 1993) and so this feature of Ku may be responsible for the observed binding to relaxed-circular DNA that did contain internal nicks. Whether ScKu1 and ScKu3 binding to supercoiled DNA was affected by secondary structures, or if the mode of binding was identical to interactions with relaxed-closed DNA, it is apparent these Ku do not require free DNA ends. Therefore it is of considerable interest to identify how Ku can bind to DNA ends are not available. If alternative models exist for Ku binding, it is feasible that ScKu2 and MtKu may lack this activity and so precipitated less closed DNAs during the PGB assays.

4.5 The Prokaryotic Ku Influence on *in vitro* Ligation.

The recognition and recruitment activities of prokaryotic Ku are accepted to be integral to the NHEJ process during repair of DSBs, however a direct role in the facilitation of ligation remains unclear. Although Ku may be required to promote location of an appropriate DNA ligase to DSBs, there have been conflicting reports regarding formation of the synapsis between DNA termini prior to ligation (Cary et al., 1997; Ramsden and Gellert, 1998). In eukaryotic NHEJ, it is hypothesised that DNA termini synapsis is regulated by both Ku70/80 and DNA-PKcs, but, as prokaryotes lack the DNA-PKcs, a synapse must be formed independently of this protein. As described in Section 4.3.2, competition *in vitro* SMB DNA binding assays investigating DNA binding indicated ScKu was able to interact simultaneously with free DNA molecules and DNA molecules attached to SMBs. To investigate further if prokaryotic Ku can facilitate ligation either by aligning DNA termini or by an alternative manner, the joining of DNA ends by T4 DNA ligase was analysed with or without the inclusion of prokaryotic Ku during ligation reactions. This DNA ligase was chosen as it was not expected to promote any specific protein-protein interactions with the Ku proteins of interest.

4.5.1 T4 DNA Ligase Ligation.

Prior to exploring the effect of the inclusion of Ku proteins during ligation reactions, it was necessary to appreciate T4 DNA ligase activity independent of Ku. Initially, pUC18 plasmids

were prepared and subjected to restriction digestion to produce linear DNA molecules containing either overhang-ends (digested with *Eco*R1) or blunt-ends (digested with *Sma*I). T4 DNA ligase activity is known to be strongly affected by the DNA substrate and ligation of DNA termini with overhang-ends is considerably more efficient than ligation of linear DNA with blunt-ends (Tait et al., 1980).

Ligation of linear plasmids can result in closed-circular DNA (ccDNA) that exhibit migration characteristics that are highly dependent on the presence of intercalators such as ethidium bromide (EtdBr) (Bowater, 2005). One molecule of EtdBr can insert between 2 base pairs of a dsDNA helix and increase the distance between these bases, causing local unwinding of the helix and an increase in writhe. This means that molecules that are initially negatively supercoiled become increasingly relaxed and eventually positively supercoiled in the presence of EtdBr. For relaxed ccDNA, the increase in writhe induces positive supercoiling and this causes these DNAs to migrate further during electrophoresis than linear DNAs of equal size. This occurs as linear molecules have no writhe and so the intercalation of EtdBr has relatively little effect on their migration during gel electrophoresis. Therefore, during gel electrophoresis in the presence of EtdBr, ligated pUC18 that is converted into closed circles will migrate further than linear pUC18 (Figure 4.14).

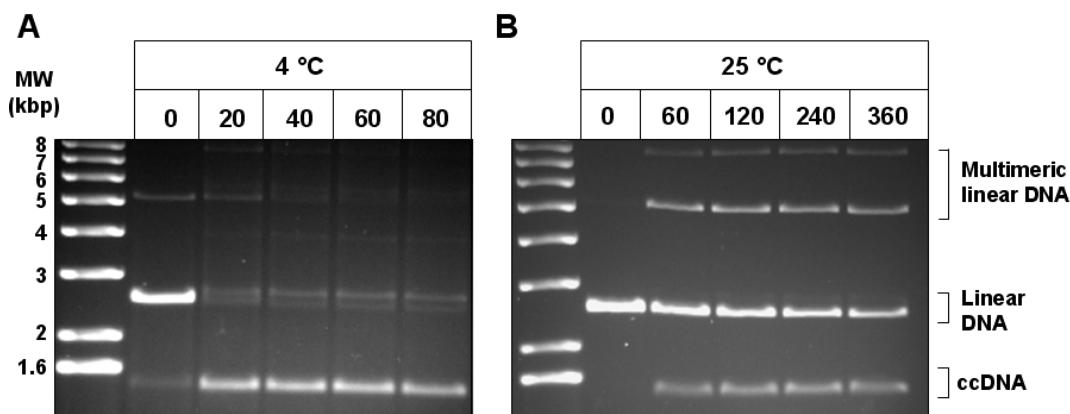


Figure 4.14 Ligation of linear pUC18 with overhang or blunt-ends by T4 DNA ligase. Ligation of 200 ng of linear DNA was performed at either 4 °C or 25 °C for varying lengths of time (shown in minutes) depending on the rate of reaction. Ligation products were analysed by electrophoresis on 0.8% agarose gels that contained EtdBr. **A.** Ligation of linear pUC18 with overhang-ends (created by *Eco*R1 digestion). **B.** Ligation of linear pUC18 with blunt-ends (created by *Sma*I digestion).

As predicted, ligation of linear DNA with overhang-ends was significantly faster than blunt-end DNA. As shown in Figure 4.14 A at 4 °C after 20 minutes, almost all linear pUC18 had been ligated into covalently closed relaxed DNA, as indicated by bands exhibiting highest electrophoretic motility. In contrast, ligation of blunt-end pUC18 (Figure 4.14 B) was

considerably slower and, even at a higher temperature of 25 °C, less ligated product was formed after 6 hours than ligation of overhang ends after only 20 minutes. Interestingly, blunt-end pUC18 appeared more susceptible to forming multimer complexes as indicated by bands at approximately 5.4 kbp (dimers) and 8 kbp (trimers) alongside covalently closed relaxed DNA.

4.5.2 Determining How Prokaryotic Ku Influence Ligation by T4 DNA Ligase.

It was of interest to investigate the effect ScKu and MtKu proteins might have upon T4 DNA ligation of both blunt and overhang ended linear pUC18. It was anticipated that the presence of Ku on DNA termini may influence the rate of ligation, either positively by assisting the alignment of DNA termini, or negatively by preventing efficient T4 DNA ligase activity. Initially, reactions were performed on blunt-end pUC18 with or without various amounts of prokaryotic Ku or BSA. It was discovered that without Ku, ligation by T4 DNA ligase was identical to that shown in Figure 4.14 B, however the addition of ScKu1 or ScKu2 showed dramatic differences as ligation products could not be seen and a large proportion of pUC18 remained trapped in the gel wells (data not shown). The inclusion of ScKu3 also demonstrated differences as several other DNA species were detected that had significantly higher molecular weights than 6 kbp. In contrast, MtKu failed to produce any ligated product (as previously found (Weller et al., 2002)) or cause DNA to remain in wells. The amount of linear pUC18 present in these reactions did appear to decrease over time to an extent, however this may result from other factors including contaminating DNase proteins. BSA was investigated as a control to identify if differences detected with Ku proteins were dependent on molecular crowding. As results from reactions including BSA were identical to reactions without any additional proteins, it was accepted that the results observed with Ku present were indeed dependent on Ku activity.

Additional experiments found pUC18 remaining in the wells was actually caused by bound protein that could be removed by treating samples with 1% SDS and heating at 65 °C that released DNA (data not shown). Hence, all subsequent assays investigating the effect of prokaryotic Ku on T4 DNA ligase activity were performed with this additional step.

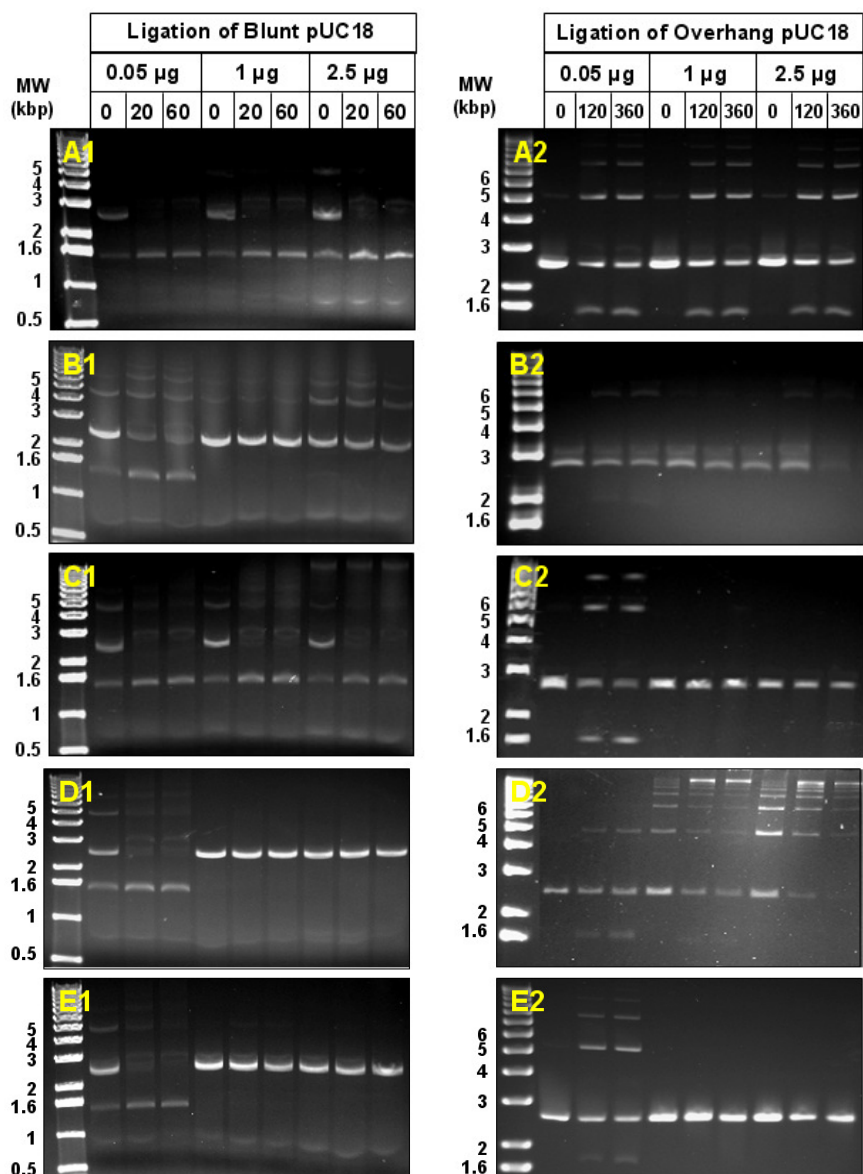


Figure 4.15 Ligation of blunt-end and overhang-end linear pUC18 by T4 DNA ligase with the inclusion of various amounts of prokaryotic Ku or BSA. Ligation of 200 ng of linear DNA was performed at 4 °C (over-hang) or 25 °C (blunt) for varying lengths of time (shown in minutes) and containing amounts of protein ranging from 0.05 - 2.5 µg in every 20 µl reaction. Ligations were stopped by the addition of appropriate amounts of 6 x native loading buffer and stored at 4 °C. DNA was analysed on 0.8% agarose gels containing EtdBr. **A1** – Effect of BSA on blunt-end ligation. **A2** – Effect of BSA on overhang-end ligation. **B1** – Effect of ScKu1 on blunt-end ligation. **B2** – Effect of ScKu1 on overhang-end ligation. **C1** – Effect of ScKu2 on blunt-end ligation. **C2** – Effect of ScKu2 on overhang-end ligation. **D1** – Effect of ScKu3 on blunt-end ligation. **D2** – Effect of ScKu3 on overhang-end ligation. **E1** – Effect of MtKu on blunt-end ligation. **E2** – Effect of MtKu on overhang-end ligation.

As evident in Figure 4.15, ligation by T4 DNA ligase in the presence of prokaryotic Ku proteins was greatly affected and in most assays, Ku appeared to completely inhibit ligation or induce the formation of multimer DNAs. When compared to Figure 4.14 where ligation

reactions were performed without additional proteins, it is apparent that ScKu1 prevents circularisation of overhang-linear DNA and completely prevents ligation of blunt-end DNA. It is noticeable in Figure 4.15 B1 that multimers of pUC18 can still be formed when high amounts of ScKu1 are present, even though circularisation is prevented, suggesting ScKu1 may facilitate joining of separate individual DNAs over joining of the same DNA. ScKu2 also appeared to influence ligation as although circularisation of overhang-end DNA was still occurring, DNA molecules of very high molecular weights were apparent when 2.5 µg of ScKu2 was present. It is likely these bands represent species consisting of multiple pUC18 molecules that have been ligated together to form large multimers. ScKu2 did negatively affect ligation of blunt-end pUC18 as very little ligatable product was detected. The inclusion of ScKu3 displayed variations on activity seen by ScKu1 and ScKu2 as circularisation was strongly inhibited when present for both pUC18 molecules, however, the facilitation of multimer formation of blunt-end pUC18 was evident by bands with molecular weights above 4 kbp (Figure 4.15 D2) that become more apparent as the amount of ScKu3 increases. It is also noticeable that, as time progresses, the sizes of the multimers also increases, strongly indicating that the T4 DNA ligase is joining together DNA termini of different pUC18 plasmids, rather than forming ccDNA as shown when no Ku proteins are present. Figure 4.15 E1 and E2 show that the introduction of MtKu prevents formation of any ligated product, showing that the inhibitory effect exhibited by ScKu proteins is not unique to *S. coelicolor* Ku proteins. To confirm the observed findings were a direct result of Ku activity and not a consequence of unspecific protein affecting reactions, for example by molecular crowding, BSA was also investigated. The results in Figure 4.15 A1 and A2 were almost identical to assays that did not include additional proteins and none of the effects identified in the presence of Ku were observed. This implies the circularisation inhibition activity is Ku-dependent and may occur because Ku situated on the DNA ends is preventing T4 DNA ligase access to DNA termini.

4.6 Discussion.

Eukaryotic NHEJ and its components have been extensively investigated (Critchlow and Jackson, 1998; Weterings and Chen, 2008). The identification of NHEJ protein homologues in prokaryotes reversed general opinion that this process was restricted to eukaryotes. Although investigations into prokaryotic NHEJ have suggested that this pathway is a two-component system comprising of a homodimeric Ku and a multi-functional ligase (Della et al., 2004), *in silico* studies on the prokaryotic *S. coelicolor* genome revealed the existence of 3 separate genes homologous to eukaryotic Ku70/80 and also other genes homologous to proteins implicated in eukaryotic NHEJ (Sayer, 2008). These findings imply that the

currently accepted two-component system of prokaryotic NHEJ may not apply to all prokaryotic organisms.

Three Ku genes from *S. coelicolor* and 1 Ku gene from *M. tuberculosis* were successfully over-expressed in *E. coli* via the pET-expression system and purified by affinity chromatography using the presence of polyhistidine tags present on the resultant recombinant proteins. Electrophoretic analysis of purified proteins revealed that migration of all prokaryotic Ku proteins was considerably lower than was expected, creating a degree of uncertainty regarding the actual identities of the proteins. Western blot analysis using antibodies specific for polyhistidine tracts confirmed that the bands observed using Coomassie staining were likely to be the recombinant Ku. However, to conclusively ascertain the identity of the recombinant proteins, ScKu1, ScKu2 and ScKu3 were subject to MALDI-TOF peptide mass fingerprint analysis. The results collected from this analysis confirmed with a high degree of confidence that the recombinant proteins were ScKu proteins derived from the relevant genes and that the observed reduction in electrophoretic mobility was caused by unknown factors. It was postulated that higher than average isoelectric points of the prokaryotic Ku proteins were influencing migration by affecting the overall charge of the proteins, meaning that migration of Ku was not relative to the migration of proteins that constituted the marker.

Once analysis had confirmed the relevant recombinant Ku proteins had been attained, DNA binding assays were performed using 2 separate *in vitro* DNA binding techniques that both encompass the precipitation of protein-DNA complexes by magnetic bead separation. As both eukaryotic and prokaryotic Ku are recognised as having high binding affinity towards dsDNA and comparatively low affinity towards ssDNA, initial assays concentrated on comparing protein binding to short dsDNA and ssDNA substrates. It was found that the prokaryotic Ku proteins demonstrated considerable differences in DNA binding activity towards both DNA substrates. Both ScKu2 and MtKu exhibited activity that has become a hallmark of Ku-DNA binding, whereby binding to dsDNA is significantly higher than to ssDNA. As it has previously been documented that MtKu binds preferentially to dsDNA over ssDNA using EMSA (Weller et al., 2002), the agreement shown during this study indicates the reliability and reproducibility of the SMB assay. Interestingly, ScKu1 had alternative activity, and binding was consistently high to either DNA. Due to the strength of this interaction even at high ionic strength, this activity is unlikely to be non-specific, and may indicate an alternative role for ScKu1 that other Ku are unable to perform. ScKu3 binding to short oligonucleotides was comparatively lower than all other proteins studied, however ensuing PGB assays did detect binding to plasmid DNAs, proving that this recombinant

protein was able to bind to certain DNA molecules. Potentially, ScKu3 binding is more DNA length dependent than other Ku and requires larger DNA molecules to initiate binding, a factor that is quite contradictory considering ScKu3 is the smallest prokaryotic Ku studied. Another cryptic finding from SMB DNA binding analysis was that ScLigD consistently displayed very selective DNA binding properties as binding to ssDNA was very efficient but binding to dsDNA was non-existent. Previously, it has been found ScLigD can ligate 40 base pair dsDNA containing an internal nick (Sayer, 2008), albeit at a relatively low rate, indicating ScLigD can at least bind to nicked dsDNA. To further quantify the observed differences in binding activity of prokaryotic Ku, affinity investigations were employed in order to attain apparent K_d values that would highlight the apparent disparity in DNA binding activity between ScKu1, ScKu2 and MtKu towards short dsDNA and ssDNA molecules. When single-protein binding was assumed, dissociation constant values for ScKu1 and ScKu2 towards dsDNA were not calculable as high concentrations of protein bound even when very low concentrations of DNA were available. One factor that may contribute to this observed activity is that multiple proteins may bind to single DNA molecules and this would contribute to the very high protein binding exhibited even at low concentrations of DNA. As Figure 4.16 shows, if 1 protein binds to 1 DNA molecule, more DNA is required to saturate the protein than if 2 proteins can bind to a single DNA. This effect is again increased if a protein binds as a tetramer as protein will very quickly become saturated and the addition of more DNA becomes ineffectual.

Irrespective of the difficulties in obtaining K_d values for all proteins studied, the values that were acquired did correlate with experiments investigating protein percentage binding to dsDNA and ssDNA. The specificity in ScKu2 binding to dsDNA and ssDNA is again evident, however a preference towards dsDNA over ssDNA was also noticeable for ScKu1. This may suggest that binding activities do not differ to the extent shown during percentage binding experiments as both ScKu1 and ScKu2 may prefer dsDNA, however ScKu1 can bind with higher affinity towards ssDNA but this difference was not detected using original conditions. The MtKu K_d values were higher than both ScKu proteins, which again agreed with percentage binding experiments that identified MtKu as having the lowest binding activity excluding ScKu3.

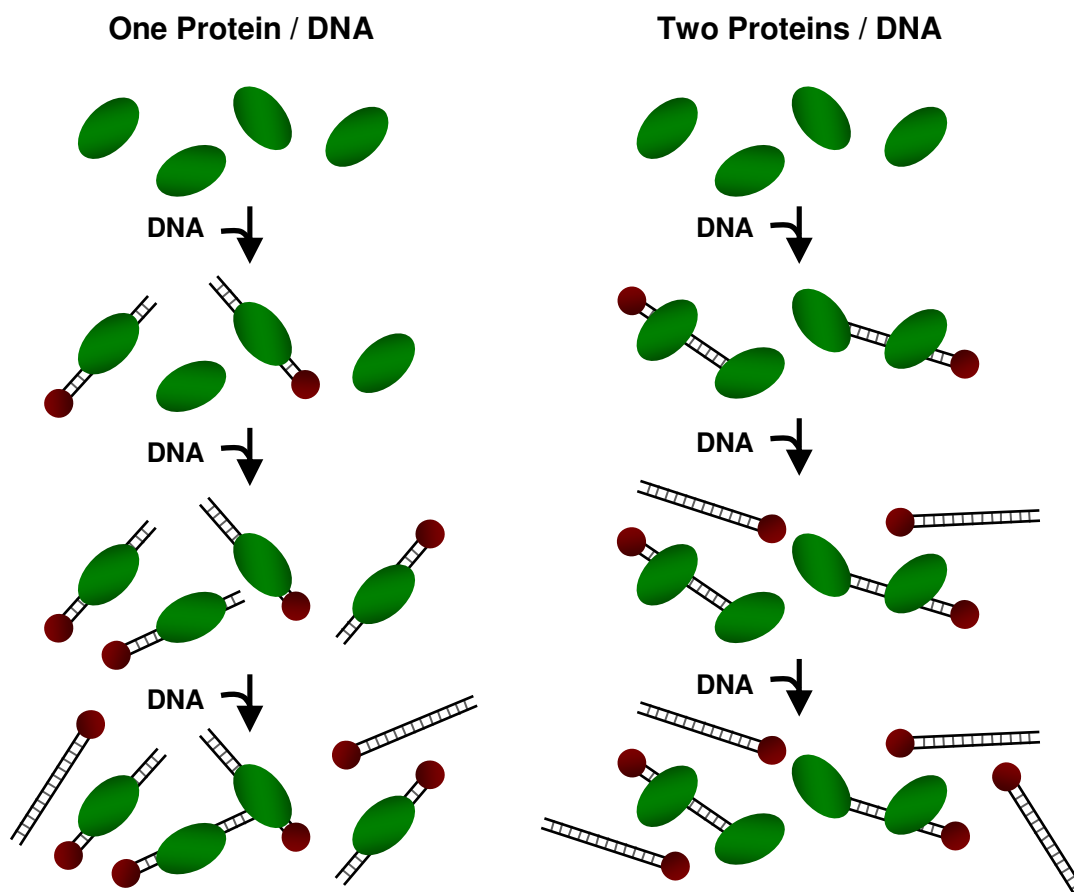


Figure 4.16 The effect of single-protein or multiple-protein binding to DNA. The mode of protein-DNA interaction will heavily influence K_d values obtained from binding affinity experiments. If 1 protein molecule binds to 1 DNA molecule, a higher amount of DNA will be required to saturate protein compared to a protein that can bind multiple molecules to a single DNA.

Competition assays provided unexpected findings that could have considerable implications in the specific cellular activities of ScKu1 and ScKu2. Initial competition assays focused on the ability of competing DNA molecules to remove Ku proteins that had already formed complexes with DNA attached to SMB (DNA-SMB). It was ascertained that the addition of free DNA post Ku incubation with DNA-SMB had negligible effects, even when free DNA was added in excess of DNA-SMB. It was possible that Ku could not be removed from the initial protein-DNA complex and so the addition of extra DNA was ineffectual. To examine this possibility, free DNAs were added simultaneously to DNA-SMB and Ku proteins so that both free DNA and DNA-SMB had equal opportunities for Ku binding. Again, it was concluded from the results that Ku was still being precipitated along with DNA-SMB, even when free DNAs were added 3 times in excess of DNA-SMB. Two likely possibilities that could explain the observed discrimination by ScKu were that either the DNA-SMB is conformationally different to free DNA as a consequence of the close association with a large magnetic bead, or that Ku were interacting with DNA-SMB and free DNAs simultaneously,

and effectively precipitating complexes of SMB-DNA-Ku-free DNAs. A modified version of the SMB assay that used fluorescently labelled ssDNAs as competing molecules was utilised to identify if free DNA was precipitated in the presence of Ku. During this analysis, assays were performed as usual except that instead of quantifying protein binding by measuring protein in unbound or bound samples, both samples were measured for fluorescence emission that was dependent on whether free DNA had been precipitated or not. The results indicated that the inclusion of ScKu1 and, to an extent, ScKu2 did influence precipitation of free DNA. This was observed by both a decrease in fluorescence in ‘removed free ssDNA’ and a slight increase in fluorescence emission of samples containing SMB-DNA-Ku. It was believed ScKu1 had a more considerable effect on the precipitation of fluorescent ssDNA as ScKu2 has previously been found to have significantly less affinity towards this substrate, however, as an unbiotinylated fluorescently labelled dsDNA was not available, only ssDNA was investigated.

The juxtaposition of 2 DNA termini is an important step for efficient NHEJ. Regions of microhomology between DNA termini may facilitate alignment to an extent, however it is thought specific protein activity is crucial for this process. Independent eukaryotic NHEJ studies have produced conflicting reports that either suggest Ku70/80 can form a synaptic complex of DNAs (Cary et al., 1997; Ramsden and Gellert, 1998) or that DNA-PKcs is required for this process (DeFazio et al., 2002; Lees-Miller and Meek, 2003). Although it may remain unclear how end-alignment occurs in eukaryotes, end-alignment in prokaryotes must be DNA-PKcs independent. The polymerase domain of *M. tuberculosis* LigD contains a functionally analogous structural element thought to facilitate connectivity between discontinuous DNA ends (Pitcher et al., 2007), indicating this protein may participate in formation of a synaptic complex, potentially alongside Ku. The SMB competition assays and subsequent fluorescence analysis during this study indicates ScKu proteins may be capable of synapse formation in *S. coelicolor*. The results from fluorescence analysis may not justify the competition assay results, however, there are potential reasons for this. Unlike free DNAs with 2 free-ends, as fluorescent ssDNAs have a fluorophore at one end, only one end will be available for alignment. Conversely, unlabelled free DNAs could potentially form long complexes consisting of multiple DNAs that may be precipitated (Figure 4.17 A). As labelled ssDNAs can only form a synapse consisting of 2 DNAs, the opportunity for precipitation is significantly lower (Figure 4.17 B). There is also the possibility of 2 fluorescently-tagged ssDNAs forming a complex together that would not be precipitated. Finally, as alignment of dsDNAs may be more efficient than ssDNAs (especially for ScKu2), this would explain why less fluorescently-labelled DNA was precipitated than anticipated.

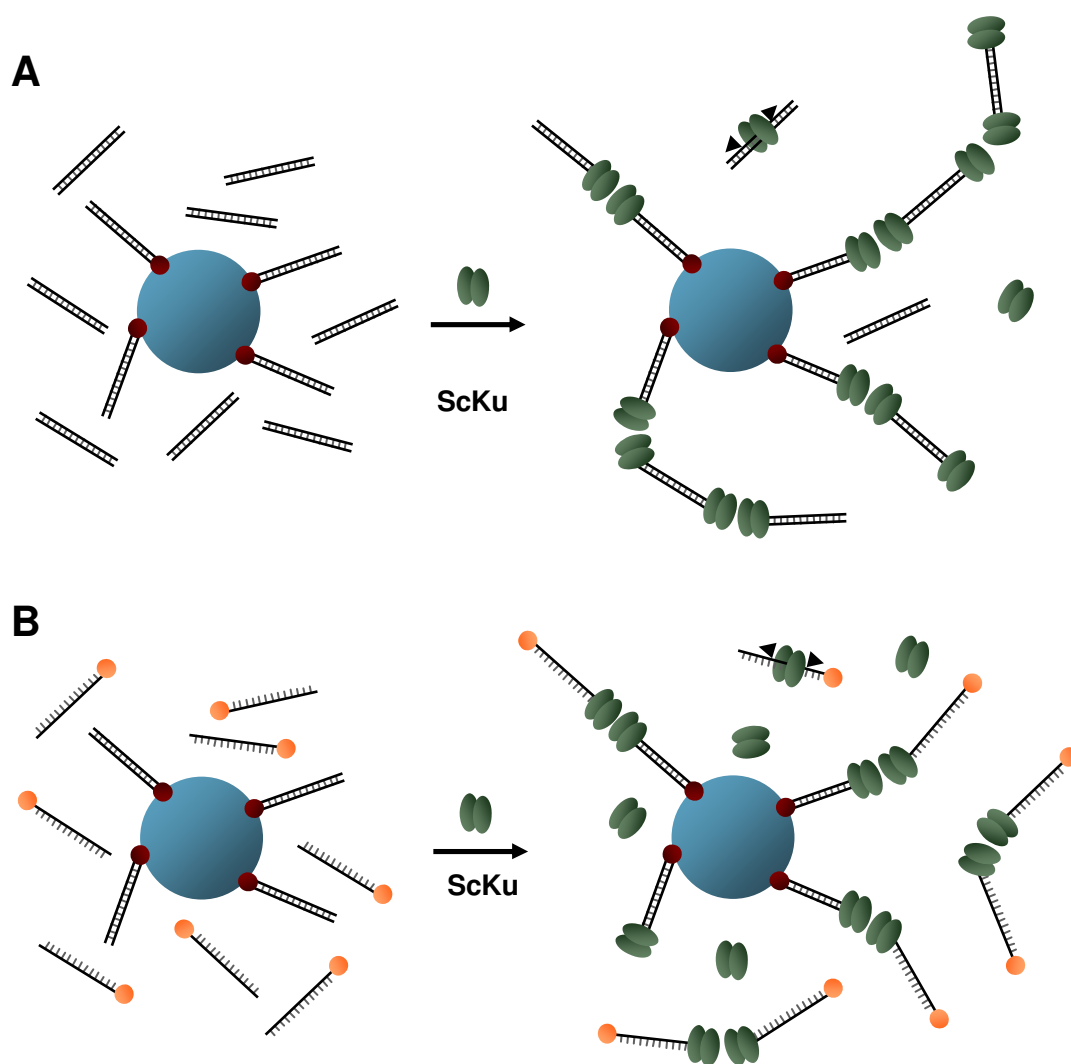


Figure 4.17 Analysis of synaptic-complex formation by ScKu1 and ScKu2 using a modified SMB DNA binding assays. To understand if the negligible effect of adding free DNAs to SMB reactions was a consequence of precipitation of free DNAs by ScKu mediated synapse formation, fluorescently tagged free DNA precipitation was measured in the presence of ScKu. **A.** Graphic representation of synaptic complex formation by ScKu. Multimers of DNAs are able to form as free DNAs possess 2 free-ends. ScKu may also be able to dissociate from free DNAs following translocation. **B.** SMB competition assays using fluorescently labelled ssDNAs – only a single ssDNA can be aligned with DNA-SMB and dimers of 2 labelled ssDNAs may also arise if ScKu is implicated in DNA alignment.

Although the proposed alignment activity by ScKu1 and ScKu2 is an attractive model that would justify how synapsis of DNA ends could occur in prokaryotic NHEJ without the requirement of eukaryotic DNA-PKcs, other possibilities remain that could also explain the conformational differences between DNA-SMB and free DNA. The ability of Ku to translocate along DNA has for a long time been a recognised activity of eukaryotic Ku (Paillard and Strauss, 1991). Considering prokaryotic Ku may also translocate along DNA, supposedly in search of threatening DSBs, it is possible ScKu proteins were binding to free DNAs and effectively sliding-off due to the restrictive size of these molecules. In contrast,

DNA-SMB not only has a biotin molecule at the 5' end, but also a large magnetic bead – an obstacle a protein could not surpass. Therefore, this behaviour may also have contributed to the apparent preference towards DNA-SMB.

Prior to analysis of prokaryotic Ku binding to plasmid DNA using an alternative PGB DNA binding assay, it was quickly apparent that the addition of plasmid DNA was stimulating the rapid degradation of ScKu proteins in a protein-specific manner. Purification of plasmid DNA significantly reduced ScKu degradation, indicating that a contaminating molecule such as a protease was responsible for the previous observations. As a general consensus is that Ku would remain encircling the DNA following DSB repair and that a process such as proteolytic cleavage may be necessary for dissociation (Schar et al., 1997; Walker et al., 2001), the rapid protein-specific degradation of ScKu observed in this study may be an example of this process. Accordingly, the high susceptibility of Ku towards degradation would support the degradation-dependent dissociation model.

For characterisation of prokaryotic Ku DNA-binding activity towards larger linear DNA, relaxed-circular DNA and supercoiled DNA, an *in vitro* PGB assay was utilised that, unlike the SMB assay, is not restricted to biotinylated substrate molecules. Assays detecting ScKu binding demonstrated that ScKu were able to bind efficiently to plasmid DNA, however, similarly to SMB assay results, the 3 ScKu proteins revealed differential binding activity. At low ionic strength, both ScKu1 and ScKu3 bound very efficiently to all plasmid molecules including supercoiled DNA, however ScKu2 binding was considerably lower. As ionic strength was increased, more specific interactions became apparent and it was ascertained that ScKu1 and ScKu3 preferred linear DNA with overhang-ends over blunt-ends. Previous eukaryotic Ku studies suggest that Ku binding to DNA ends is independent of the exact structure or sequence (Falzon et al., 1993; Paillard and Strauss, 1991), meaning there should be no difference between blunt or overhang-ends. The ability for ScKu1 and ScKu3 to discriminate between blunt and overhang-ends is only apparent at higher KCl concentrations and may explain why previous investigations have not recognised this specific activity. Whether the preference towards linear-overhang DNA is dependent on the structure or differences in sequence at the DNA termini compared to linear-blunt DNA remains uncertain.

Interestingly, ScKu3 demonstrated stronger interactions with supercoiled DNA over linear blunt-end DNA, an unusual result considering Ku do not ordinarily interact with closed-DNAs (Dylan and Yoo, 1998). The absence of available DNA ends that are believed to be important in Ku binding indicate that Ku may be capable of interacting with DNA in an alternative mode other than DNA end recognition. Potentially, Ku may directly recognise

secondary structures that can arise in closed DNA molecules and these structures could thread through a Ku dimer channel in a similar fashion to dsDNA ends. Consistent with this idea, previous studies also identified Ku could interact *in vitro* with DNA independent of the availability of DNA ends and found certain structural features, including hairpins could induce Ku binding to the same extent as dsDNA with free ends (Paillard and Strauss, 1991). Studies have also shown that Ku can interact with circular molecules containing an internal nick (Blier et al., 1993) with identical affinity to dsDNA with free DNA ends, and also to relaxed ccDNA (but with lower affinity). It would not be possible for Ku to bind to these circular DNA molecules in the currently accepted model as there is no available DNA end or hairpin-structures to thread through the central Ku channel. Therefore, it is possible that Ku may exhibit multiple modes of DNA binding, such as dimerisation following interaction with DNA to allow binding to closed molecules and that this alternate mode of binding was required for the observed interactions with supercoiled pUC18 during the PGB assay.

To discover if the observed binding to ccDNA was dependent on secondary structures, or if ScKu could still interact with circular DNA exhibiting reduced secondary features, relaxed-circular DNAs were produced by restriction enzyme digestion resulting in a single nick that releases the stored energy of the supercoiled molecule. The data from PGB assays using relaxed-circular DNA as prokaryotic Ku substrate found almost identical results as for supercoiled DNA, indicating that secondary structures may not have been the only determinant for Ku-supercoiled DNA binding. As an internal nick in circular-DNA has previously been shown to initiate eukaryotic Ku binding (Blier et al., 1993), this feature may have influenced the observations during this experiment, however, it still remains unknown how Ku can bind to such DNAs if no free end or unusual structure is present.

An unexpected observation was that ScKu2 consistently exhibited the least plasmid-DNA binding activity out of the 3 ScKu proteins, and ScKu3 (which had shown relatively low activity during SMB assays) actually bound efficiently to several plasmid DNAs. Due to the flexibility of the PGB assay, identical substrates used during SMB assays could be used as substrates instead of plasmid DNA. This provided an opportunity to assess agreement between both assays and it was determined that there was very high correlation between results, except that ScKu3 appeared to precipitate more ssDNA than previously detected during SMB analysis. The comparatively low ScKu2 binding towards plasmid DNA was not anticipated, however potential interference from factors present in the PGB assay, such as antibodies, could affect protein activity. Conversely, these factors may actually enhance ScKu3 activity and, thereby, increase DNA binding.

The PGB assay had successfully detected ScKu binding to both plasmid and oligonucleotide substrate and it was of interest to elucidate if MtKu had comparable activity. Under identical conditions to ScKu assays, MtKu bound very poorly to any substrate DNA, however, when temperature was increased, MtKu binding significantly improved from less than 10% to approximately 90% bound to linear overhang DNA. It was recognised that in comparison to ScKu that demonstrated activity over a wide temperature-range, MtKu is only efficient at higher temperatures. These findings suggest that the MtKu protein is better adapted to the typical conditions that *M. tuberculosis* would be expected to survive than at lower temperatures. Though the results suggested that the optimum temperature for activity was around 25 °C and not the anticipated 37 °C, other factors would need to be considered that would affect activity such as stabilising components that would be absent during *in vitro* assays. Interestingly, the results also revealed that, similarly to ScKu, MtKu showed a preference towards overhang-ends compared to blunt-ends at all temperatures investigated. Although previous studies have found Ku binds to DNA ends independent of sequence or conformation (Falzon et al., 1993), the data from this study suggests prokaryotic Ku have a preference towards ends with overhangs, though this difference may not be detectable in less sensitive assays. Although MtKu demonstrated the apparent preference towards linear DNA with overhang-ends as identified with ScKu1 and ScKu3, binding to supercoiled DNA was significantly lower and, therefore, exhibits the currently accepted view regarding Ku DNA binding activity.

The recruitment of a multifunctional DNA ligase during prokaryotic NHEJ is currently accepted to be an important function of Ku. Accordingly, it was speculated that Ku may influence ligation reactions either by direct interactions with a DNA ligase protein, or by modulating the DNA into a more favourable substrate. Assays were developed that monitored T4 DNA ligase joining of linear plasmid DNA with or without prokaryotic Ku or BSA proteins included in reactions. Initial observations of T4 DNA ligase activity illustrated that ligation of overhang-end DNA was considerably more efficient than blunt-end ligation. Blunt DNA also appeared to be more prone to forming multimer molecules alongside circularisation. Data from assays that included prokaryotic Ku revealed that, in most instances, ligation of both types of linear DNA was completely inhibited when high amounts of Ku were present. This inhibitory effect may in part result from the presence of Ku proteins located on DNA termini that directly obstruct ligation, meaning that even if prokaryotic Ku are able to translocate along DNA, some protein also remains at the ends of the DNA molecule, presumably to interact with other NHEJ components. It was also apparent that although Ku were preventing circularisation of plasmid DNA, the formation of multimeric DNA was evident, particularly with ScKu3. This activity has been previously documented

(Paillard and Strauss, 1991) with eukaryotic Ku and this occurrence was proposed to arise following rigidification of DNA by the presence of several internally bound Ku that inhibit circularisation by reducing flexibility.

Similar findings were also observed in 2 earlier independent studies investigating how Ku affects T4 DNA ligase ligation. In one study it was found that human Ku was highly effective in inhibiting the conversion of linear-blunt end DNA into covalently closed molecules following T4 DNA ligase incubation, also leading to the conclusion that Ku situated on the DNA termini interfered with the function of the ligase (Blier et al., 1993). In a separate study investigating the effect that Ku from monkey CV1 cells has on T4 DNA ligase activity, it was found that circularisation of plasmid DNA was prevented when Ku was present and that the formation of linear oligomers was favoured over circularisation (Paillard and Strauss, 1991). It was suggested that Ku proteins were binding to plasmid DNA, and subsequent translocation along DNA allowed multiple Ku molecules to bind and effectively make the DNA more rigid, thereby preventing circularisation.

Results from this study appear to show high correlations with previous studies as prokaryotic Ku considerably reduced T4 DNA ligase ligation. Although the formation of circular DNAs was almost completely abolished when high amounts of Ku were present, in some instances it was apparent that multimer formation was still possible. This was most noticeable with ScKu3 and blunt-end pUC18, however it is unclear why ScKu3 would be most efficient at promoting multimer formation. Potentially, as this protein is the smallest of the prokaryotic Ku studied, it may obstruct T4 DNA ligase access to a lesser extent. The ability of prokaryotic Ku to protect DNA ends from degradation has recently been documented (Zhu and Shuman, 2009). It was proposed that although some Ku molecules translocate inwards from the DNA termini, the Ku protein that remain at the DSB ends prevent exonuclease mediated digestion of the DNA. It is very possible the ligation inhibition identified in this study is related to this end-protection activity and that Ku located at DNA ends are indeed preventing T4 DNA ligase access.

CHAPTER 5

Macromolecular Interactions Involving ScKu

Proteins

In vitro protein analysis of ScKu proteins implied that although all 3 display DNA binding activity that is largely characteristic of Ku proteins, clear disparity existed in their behaviour. There is little doubt that *in vitro* experiments can be hugely beneficial in the characterisation of protein function, however the ability to correlate *in vitro* data with information attained during *in vivo* analysis greatly enhances relevance to cellular function. Accordingly, it was decided to augment data collected from *in vitro* assays with experiments that explored ScKu activity directly from *S. coelicolor*. It was proposed that the cloning of recombinant polyhistidine-tagged *scku* into wild-type and knockout strains of *S. coelicolor* would be advantageous for the analysis of *in vivo* expression, complementation and macromolecular interactions. The production of antibodies specific for individual ScKu was also thought be beneficial for identification of similar *in vivo* characteristics but would also provide the opportunity for *in vitro* co-immunoprecipitation (Co-IP) assays.

Although recombinant versions of *scku1* and *scku2* genes had already been produced, to be sure that regulation of expression was similar to *in vivo* expression, promoter regions that controlled the expression of both genes were also required. It was hoped that these regions would be situated within 200 base pairs of the start of the gene and therefore, these regions would also be included in the final gene products. Gene constructs would also contain C-terminal polyhistidine tags that although not essential, would aid additional experiments. Resultant gene products were to be cloned into selected phage-derived vectors that integrate at specific locations in the *S. coelicolor* chromosome.

Finally, *in vitro* analysis of macromolecular interactions involving ScKu were performed. The high-specificity of the anti-ScKu immunoglobulins meant that Co-IP assays could be developed to ascertain if ScKu proteins function independently, or if interactions between different ScKu proteins could be detected. Assays also looked at whether ScKu proteins could form complexes with ScLigD or if the presence of ScKu bound to DNA could enhance the precipitation of other *S. coelicolor* NHEJ proteins.

5.1 Cloning of Recombinant *sku1* and *sku2* into Integration Phage Vectors.

Although his-tagged recombinant versions of *sku* genes had previously been produced that were currently sited in pET28(a) vectors (Sayer, 2008), for analysis of accurate *in vivo* expression properties recombinant *sku* genes that also contained promoter regions would be more appropriate. To achieve this, primers were designed (Table 5.1) that would amplify specific regions on the *S. coelicolor* chromosome containing either *sku1* or *sku2* and approximately 200 bp of upstream DNA that was predicted to contain promoter regions (Figure 5.1).

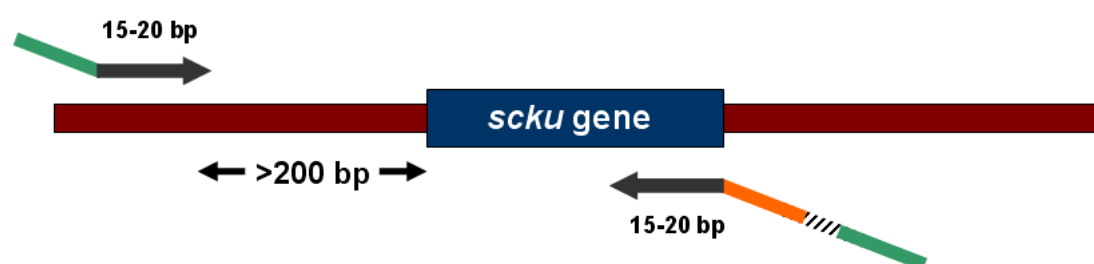


Figure 5.1 PCR cloning strategy for the production of recombinant *sku1* and *sku2* genes. Sequences for *sku1* and *sku2* genes including upstream regions were obtained from the StrepDB database. Primers containing appropriate restriction enzyme cut-sites were produced that would produce amplified fragments containing 1 blunt and 1 over-hang end. Downstream primers also contained sequences encoding a polyhistidine tag, a stop codon and an *EcoRV* cut-site.

Upstream primers consisted of 5' restriction enzyme cut-sites that would produce overhang-ends and approximately 15 – 20 nucleotides complementary to the *S. coelicolor* chromosome at sites over 200 base pairs upstream of the start of the relevant *sku* gene. Downstream primers were designed to anneal to the 3' end of the *sku* gene prior to the stop codon. For both *sku* genes, a 5' *EcoRV* cut-site was present adjacent to a sequence coding for a polyhistidine tag that was situated next to a stop codon and nucleotides complementary to the *sku* gene. The resulting PCR products would consist of: 5'- *Spe1* or *Xba1* cut-site – *sku* and upstream region – polyhistidine – stop – *EcoRV* cut-site -3'. Once amplified, recombinant genes were to be cloned directly into Zero Blunt® TOPO® as a holding vector to aid initial cloning, with analysis by PCR and sequencing to confirm cloning had been successful. Finally, inserts were to be removed by appropriate restriction enzyme digestion and ligated into *Streptomyces* integration vectors.

Many species in the *Streptomyces* genus are amenable to genetic manipulation via site-specific integration vectors derived from temperate phages. The integrase *int/attP* from

Streptomyces phage Φ C31 is a member of the serine recombinase family of site-specific recombinases that forms an essential component of integration vectors, including pSET152 (Kieser et al., 2000). Vectors containing *int/attP* can integrate at an *attB* or *attP* site in the *Streptomyces* chromosome using a mechanism of recombination similar to resolvase-invertase family members (Stark et al., 1992). Briefly, following recognition of recombination sites, a synaptic-complex is formed between the 2 DNAs. A recombinase then cleaves the 4 DNA strands, which are then positioned in a recombinant format prior to rejoining of the phosphodiester backbone. The Φ C31 *int/attP* locus contained within pSET152 integrates intragenically into the non-essential SCO3798 alongside pseudo-*attB* in certain *Streptomyces* species (Combes et al., 2002). Vectors have also been engineered that contain the *int/attP* locus from *Streptomyces* phage Φ BT1 (a homo-immune relative of Φ C31) that integrate into an alternative *attB* site in the SCO4848 gene of *S. coelicolor* (Gregory et al., 2003). Vectors including pMS82 are similar in design to pSET152 except that integration proceeds via the Φ BT1 *int/attP* instead of Φ C31. As both vectors integrate at different chromosomal sites, the opportunity exists to use 2 compatible vectors in the same organism.

Although *in vitro* assays had revealed interesting findings regarding ScKu protein activity, analysis of *in vivo* function would provide additional information about prokaryotic NHEJ. The introduction of recombinant versions of chromosomal-derived *scku1* and *scku2* into wild-type and knockout strains of *S. coelicolor* (that had previously been prepared (Handley, Bowater and Hutchings, unpublished data)) would provide an opportunity to assess gene expression, complementation and macromolecular *in vivo* DNA interactions. Delivery of recombinant *scku1* and *scku2* situated in integration vectors pMS82 and pSET152 could be achieved via conjugation between *E. coli* and *S. coelicolor* and site-specific integration of these genes would provide a platform for protein characterisation.

5.1.1 Restriction Digest Analysis of Integration Vectors pMS82 and pSET152.

Prior to cloning of *scku1* and *scku2* into pMS82 and pSET152, it was crucial to confirm that the integration vectors were as expected. Restriction digest reactions were performed on both vectors using a series of selected enzymes that were known to digest at specific sites following sequence analysis of vectors. Importantly, enzymes that would eventually be used to clone the recombinant *scku* genes into the integration vectors were also analysed to ensure only unique cut-sites were present.

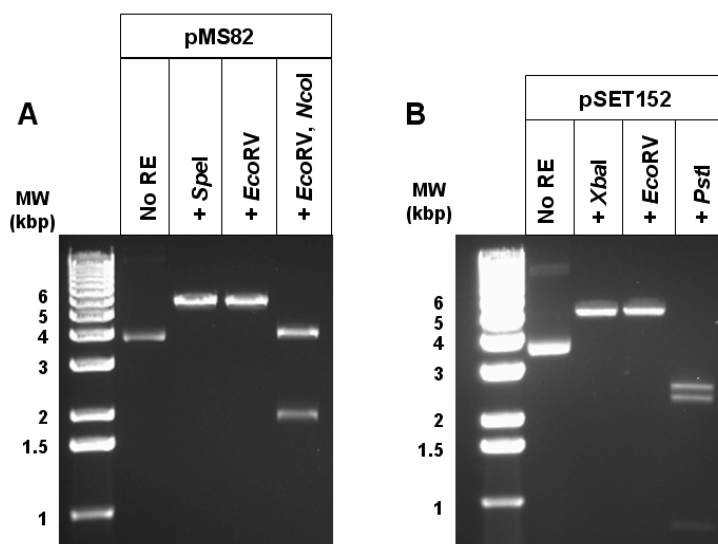


Figure 5.2 Restriction digest analysis of integration vectors pMS82 and pSET152. Approximately 200 ng of vector DNA was digested with specific enzymes under appropriate conditions for 2 hours and analysed by electrophoresis using 1% agarose gels. **A.** Analysis of pMS82 by separate digestion with *SpeI* and *EcoRV* that produce a single fragment (6108 bp) and digestion with both *EcoRV* and *NcoI* to produce 2 fragments (2036 bp and 4072 bp). **B.** Analysis of pSET152 by separate digestion with *XbaI* and *EcoRV* that produce a single fragment (5715 bp) and digestion with *PstI* that produces 3 fragments (27572 bp, 2341 bp and 802 bp).

Restriction digest analysis (Figure 5.2) confirmed that the identities of the integration vectors were as expected. Importantly, the enzymes that had been selected for future *sku* gene cloning into these vectors only cut at single locations as anticipated, meaning that the vectors and chosen restriction enzymes would be suitable for future cloning.

5.1.2 PCR Amplification of Recombinant *sku1* and *sku2*.

Preliminary PCRs focused on determining the optimum conditions required to amplify the desired genomic regions from chromosomal DNA extracted from *S. coelicolor* using Taq polymerase. The high GC-content of *Streptomyces* species can complicate PCRs and non-specific annealing of primers, formation of secondary structures in primers and primer-primer annealing can be problematic. An important component that can prevent secondary structure formation is dimethyl sulfoxide (DMSO), a polar aprotic solvent that acts by interfering with the self-complementarity of DNA (Chakrabarti and Schutt, 2001). However, at high levels DMSO can have adverse effects by increasing mutation frequency. To identify appropriate PCR conditions, reactions were performed using various concentrations of DMSO from 0 – 12% and annealing temperatures ranging from 55 – 61 °C. It was found that amplification using chromosomal DNA as a template was very inefficient for both *sku1* and *sku2* (Figure 5.3 A) as no products of the expected size were produced.

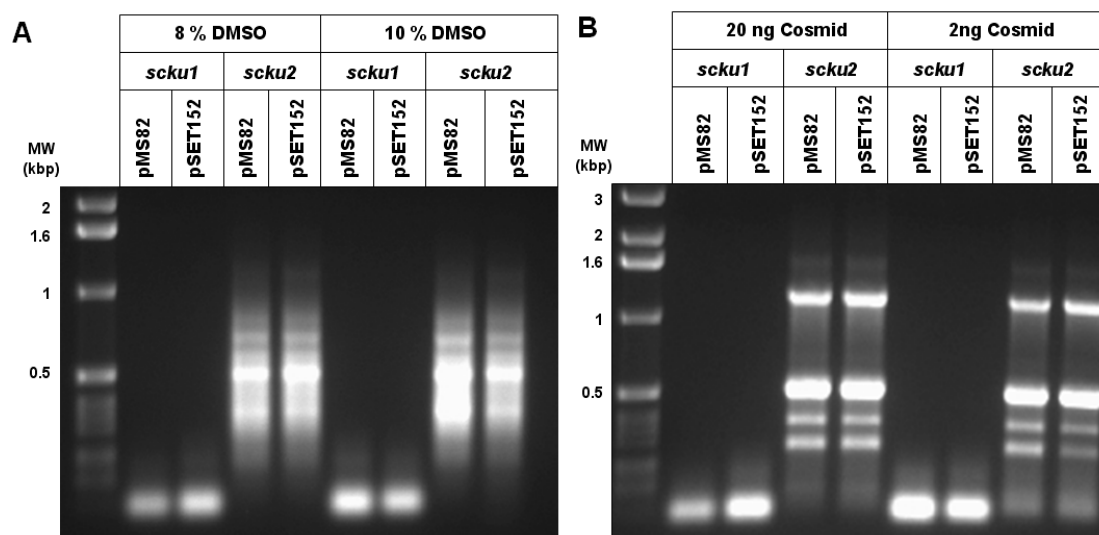


Figure 5.3 PCR amplification of *scku1* and *scku2* gene products to be inserted into pMS82 and pSET152. Conditions of PCRs by Taq polymerase were varied to allow identification of optimum reaction constituents for amplification of *scku1* (~ 1.35 kbp) and *scku2* (~ 1.3 kbp) genes. ‘pMS82’ and ‘pSET152’ refer to the primers that were used to create gene products that would eventually be cloned into pMS82 and pSET152 vectors. PCR products were analysed by electrophoresis on 1% agarose gels. **A.** PCR results using *S. coelicolor* chromosomal DNA as a template at 8% and 10% DMSO final concentrations. **B.** PCR results using 2 ng and 20 ng specific cosmid DNA as reaction templates.

Using cosmid DNA as reaction templates instead of chromosomal DNA has advantages, primarily as the ratio of specific DNA to unspecific DNA is significantly improved, thereby reducing the opportunity for unspecific primer annealing and increasing the probability of specific annealing. As with chromosomal template reactions, PCRs were performed to establish the most efficient conditions for using st6G9 (*scku1*) or st5G5 (*scku2*) cosmid DNA as the template (Figure 5.3 B). It was apparent that for *scku2*, a band present at approximately 1.3 kbp was produced following amplification with cosmid DNA that was thought to be the desired product. Increasing annealing temperature enhanced the intensity of this band and reduced the amount of other products previously detected (Figure 5.4 A). Unfortunately, *scku1* amplification was unsuccessful and there was no indication of any primer annealing. To obtain *scku2* products that were less likely to contain errors that may be incorporated during DNA amplification, PCRs were performed using a proofreading Pfu polymerase, which reduces the probability of mutations in the final PCR products. Following reactions, products at approximately 1.3 kbp were gel-extracted, purified and analysed by gel electrophoresis (Figure 5.4 B).

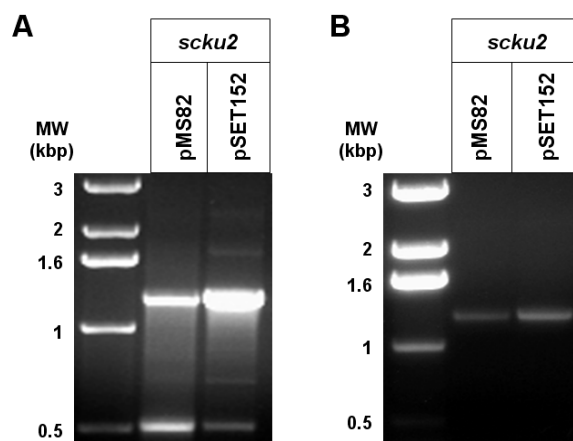


Figure 5.4 PCR amplification of *sku2* using optimum conditions and analysis of extracted and purified *sku2* product. Amplified DNA was analysed by electrophoresis using 1% agarose gels. **A.** PCR results from reactions amplifying *sku2* using Taq polymerase at an annealing temperature of 59 °C and 6% DMSO. **B.** *sku2* product gel-extracted and purified following PCR using Pfu proofreading polymerase.

Electrophoresis confirmed that the *sku2* PCR products designed to insert into pMS82 and pSET152 had both been successfully extracted and purified as stocks at approximately 30 ng / μ l. As all PCRs that were performed to amplify *sku1* had failed, it was decided at this point to focus all subsequent work on *sku2*.

5.1.3 Zero Blunt® TOPO® Cloning of Recombinant *sku2*.

Amplified *sku2* (for both pMS82 and pSET152 vectors) was initially cloned into pCR®-Blunt II- TOPO vectors. Colony PCR was performed to identify colonies that had been successfully transformed with vectors containing the *sku2* gene. It was apparent that although both *sku2*-pMS82 and *sku2*-pSET152 were similar in size and base composition, *sku2*-pMS82 appeared less apt at producing transformants. Initially, 3 colonies of *E. coli* that had been transformed with plasmids containing either *sku2*-pSET152 or *sku2*-pMS82 were analysed (Figure 5.5 A).

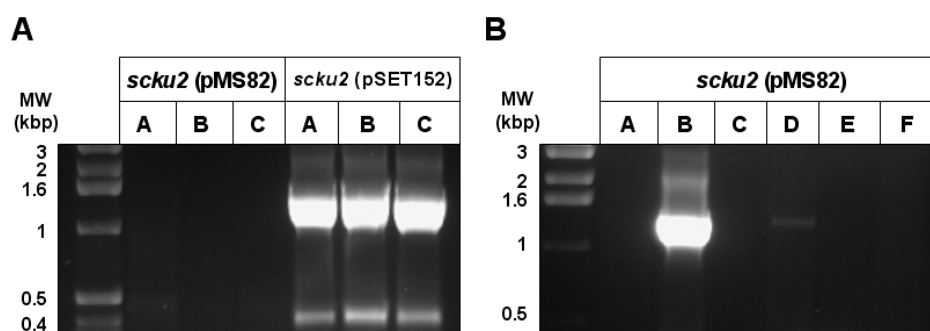


Figure 5.5 Colony PCR to detect successful transformants of *sku2*-pMS82 and *sku2*-pSET152. PCRs were performed (Section 2.2.5.3) and amplified DNA was subsequently analysed by gel electrophoresis on 1% agarose gels. Recombinant *sku2* were both ~1.3 kbp. **A.** Analysis of *E. coli* colonies that had been transformed with pCR[®]-Blunt II-TOPO vectors containing either *sku2*-pMS82 or *sku2*-pSET152. **B.** Colony PCR analysis following additional cloning of pCR[®]-Blunt II-TOPO containing *sku2*-pMS82 product.

Analysis of amplified DNA implied that the 3 colonies believed to contain *sku2*-pSET152 had been successfully transformed, however no colonies contained *sku2*-pMS82 – a probability that was verified by further DNA analysis of these colonies (data not shown). Cloning of *sku2*-pMS82 into pCR[®]-Blunt II-TOPO vectors was attempted for a second time with alternate ratios of *sku2* gene to pCR[®]-Blunt II-TOPO vector DNA and 6 of the ensuing colonies were subjected to colony PCR analysis (Figure 5.5 B). Only a single colony contained amplified product that was the expected molecular size of 1.3 kb.

Plasmid DNA was extracted and purified from colonies that had produced positive results during colony PCR and analysed by gel electrophoresis (including an additional *sku2*-pSET152 plasmid) alongside pRB240 (Figure 5.6 A). The differences in electrophoretic mobility of empty pCR[®]-Blunt II-TOPO vector compared to plasmids thought to contain *sku2* genes reinforced the probability that cloning had been successful. Digestion of cloned pCR[®]-Blunt II-TOPO vectors with *Eco*RI would theoretically release insert DNA as 2 *Eco*RI cut sites flank the region of vector that DNA can become incorporated. Analysis of DNA following this digestion again suggested the presence of *sku2* as bands were visible at approximately 1.3 kb (Figure 5.6 B). To further ascertain the presence of *sku2* in vector DNA, vectors were digested with relevant restriction enzymes that would eventually be used to sub-clone *sku2* into either pMS82 or pSET152 integration vectors.

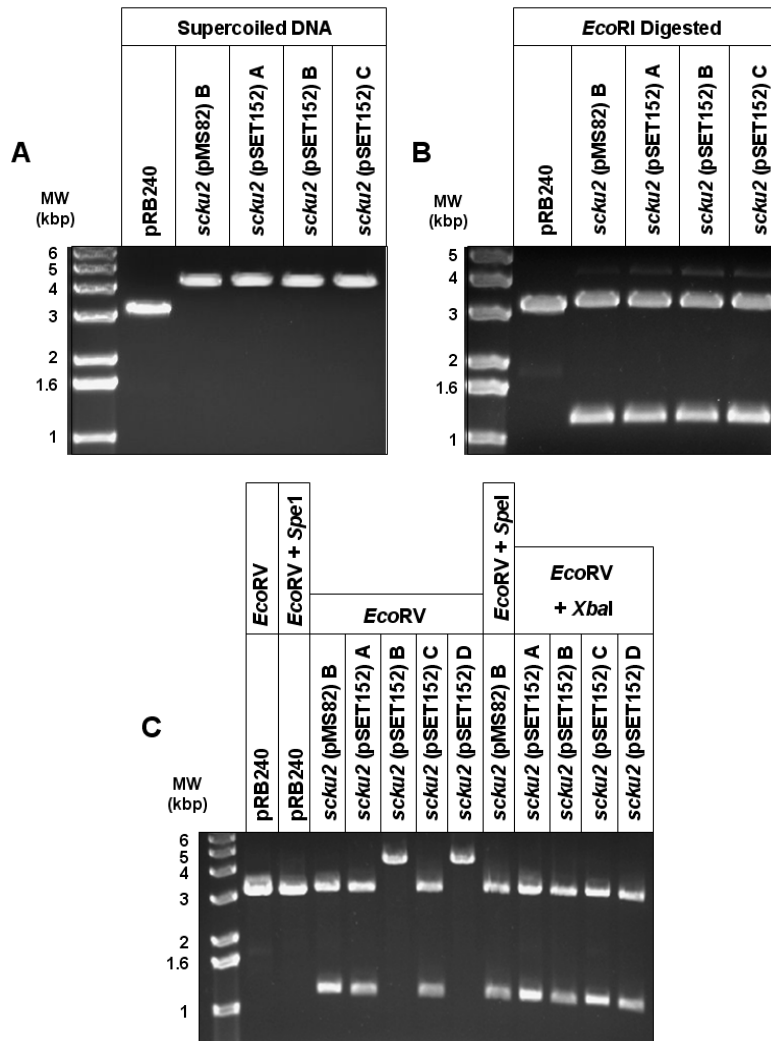


Figure 5.6 Electrophoretic analysis of DNA following cloning of *sku2* into pCR®-Blunt II-TOPO vectors. Vector DNA that did or did not contain *sku2*-pMS82 or *sku2*-pSET152 was subjected to gel electrophoresis using 1% agarose gels to determine the presence and size of insert DNA. Restriction digests using relevant enzymes confirmed the anticipated inserts were present and also determined orientation. **A.** Electrophoretic migration of supercoiled vector DNA. **B.** Electrophoretic analysis of vector DNA following *EcoRI* restriction digestion to release insert DNA. **C.** Restriction digest analysis following single or double digests with relevant enzymes. Expected fragment sizes are described in Table 5.1.

Table 5.1 Predicted fragment sizes of pCR®-Blunt II-TOPO vector DNA with or without *sku2* inserts following restriction enzyme digestion.

Vector DNA	<i>EcoRV</i>	<i>EcoRV</i> and <i>XbaI</i>	<i>EcoRV</i> and <i>SpeI</i>
pRB240	3519	-	55, 3464
pCR®-Blunt II-TOPO vector + <i>sku2</i> -pMS82	19, 4795 or 1314, 3500	19, 36, 1295, 3464	-
pCR®-Blunt II-TOPO vector + <i>sku2</i> -pSET152	19, 4795 or 1314, 3500	-	33, 1295, 3486

Single digests with *EcoRV* and double digests using *EcoRV* and either *SpeI* or *XbaI* confirmed that vectors contained the expected insert DNA. It was also apparent that the orientation of *scku2*-pSET152 varied between the 4 vector samples analysed and that vectors A and C contained insert DNA in the opposing direction to B and D. As there is no bias in orientation during blunt-end DNA cloning this occurrence was probable, however orientation would not affect future work as inserts would eventually be sub-cloned into integration vectors using specific restriction enzyme sites, meaning orientation would be controlled. Sequencing reactions were also carried out to ensure products had been obtained that did not contain mutations.

5.1.4 Cloning of *scku2* into pMS82 and pSET152.

Sub-cloning of *scku2* into both pMS82 and pSET152 integration vectors at first required the digestion of pCR[®]-Blunt II-TOPO vectors containing gene inserts and integration vectors with 2 appropriate restriction enzymes. Digested DNA was analysed by gel electrophoresis and relevant bands were then extracted and purified. Ligation of integration vectors and insert DNA were performed at molar ratios of 1:1, 1:3 and 1:6 (vector: insert) using T4 DNA ligase at 4 °C for 16 hours. Successful clones were identified by colony PCR and restriction digest analysis (Figure 5.7).

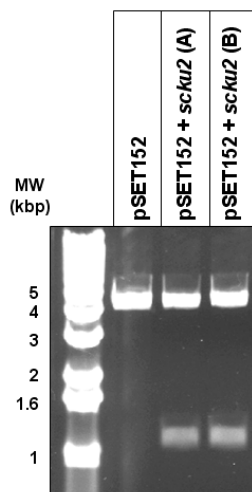


Figure 5.7 Restriction digest analysis of pSET152 containing *scku2* gene construct. pSET152+*scku2* DNA that had been obtained from 2 transformants (A and B) were digested with *XbaI* and *EcoRV* alongside pSET152 not containing *scku2*. The digested DNA was analysed by electrophoresis on 1% agarose gels. Bands at ~ 1.3 kb representing *scku2* were detected in both preparations.

Restriction digest analysis of pSET152 following sub-cloning had confirmed that *scku2* had been successfully inserted into this integration vector and this construct was now prepared for transformation into appropriate *E. coli* host strains and eventual conjugation into *S. coelicolor*. Sub-cloning of *scku2* into pMS82 had been comparatively less successful and no pMS82 containing *scku2* were identified. Although it would have been beneficial to have *scku2* in both pSET152 and pMS82, *in vivo* protein analysis would still be possible using only pSET152. For mutant *S. coelicolor* strains with apramycin resistance, a PCR-targeting approach (Section 2.2.8.4) can be used to produce pSET152+*scku2* derivatives are hygromycin resistant and can be selected following conjugation.

5.2 Scku1 and ScKu2 Immunisation and Antibody Analysis.

All previous immunodetection of ScKu described in this study was achieved using antibodies specific for polyhistidine regions present on recombinant proteins. As these antibodies do not differentiate between individual recombinant ScKu proteins, their use for immunoprecipitation of specific individual recombinant ScKu or detection of ScKu present in *S. coelicolor* would be problematic. However, the development of antibodies specific for individual ScKu proteins would be advantageous for future studies to determine various features, including protein expression and identification of protein-protein interactions. Approximately 1.5 mg of both recombinant ScKu1 and ScKu2 were purified and used for an immunisation procedure that allows for the production of polyclonal antibodies (Cambridge Research Biochemicals) specific for both individual proteins

5.2.1 Analysis of Recombinant ScKu1 and ScKu2 using Polyclonal Antibodies.

The custom polyclonal antibody procedure used during immunisation of ScKu1 and ScKu2 generates several bleeds that are obtained at various times during the process. Repeated immunisation with pure protein samples produces bleeds that contain varying amounts of polyclonal antibodies specific for individual proteins, effectively meaning that initial bleeds contain fewer specific antibodies than bleeds that are obtained following repeated immunisations. Prior to the procedure, a pre-bleed is also taken, which should not contain antibodies directed against the proteins of interest. This pre-bleed can be directly compared with subsequent test-bleeds and a final harvest bleed that should contain the highest amount of specific antibodies. For initial analysis of bleeds obtained during immunisation, known amounts of recombinant ScKu1 and ScKu2 were subject to gel electrophoresis, western blotting and immunodetection using polyclonal antibodies present in bleeds.

Due to the potential similarities of both ScKu proteins, including the polyhistidine regions present on both recombinant proteins, it was important to ensure specificity of antibodies by exposing both proteins to antibodies produced from both immunisations. As cross-reactivity of antibodies would significantly impair both *S. coelicolor* cell extract analysis and immunoprecipitation experiments, it was essential to establish optimum conditions to ensure non-specific antibody-protein interactions were at a minimum. Preliminary investigations to determine pre-immunisation bleed, primary test bleed (TB1) and test-bleed 2 (TB2) immunodetection of ScKu1 and ScKu2 are shown in Figure 5.8.

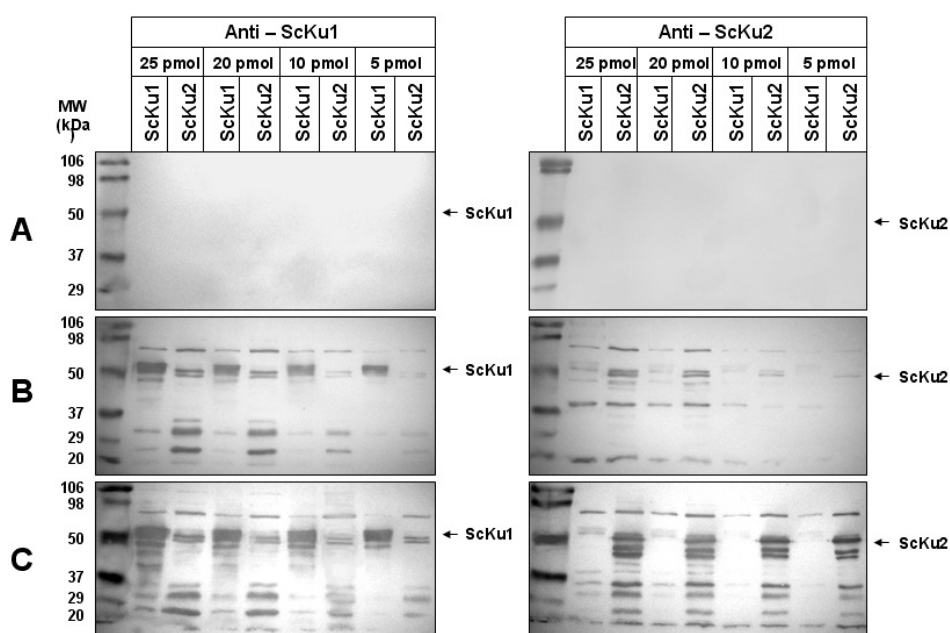


Figure 5.8 Immunodetection of ScKu1 and ScKu2 with polyclonal antibodies present in bleeds obtained from immunisation in rabbits. Varying amounts of recombinant ScKu1 and ScKu2 were subjected to SDS-PAGE and western blotting. Nitrocellulose membranes were then incubated in primary polyclonal antibody (1:10000) at 4 °C for 16 hours and then anti-rabbit secondary antibody (1:10000) for 2 hours at 22 °C. **A.** Immunodetection with pre-immunisation antibodies. **B.** Immunodetection with test-bleed 1 (TB1). **C.** Immunodetection with test-bleed 2 (TB2).

Early results from immunodetection experiments revealed that antibodies in TB1 and TB2 were successfully detecting ScKu proteins even at relatively low protein amounts of 5 pmol. As pre-bleeds from both immunisations consistently failed to detect any proteins, the observed antibody-protein binding present in TB1 and TB2 were thought to be a direct consequence of immunisation. It was also apparent that contaminating *E. coli* proteins present following purification were also detected by antibodies, however, this immunodetection was lower than specific antibody-ScKu interactions and non-specific binding was largely reduced at lower levels of protein. It was observed that antibodies were cross-reacting as when higher amounts of proteins were present, anti-ScKu1 was detecting

ScKu2 and anti-ScKu2 was detecting ScKu1 (albeit to a much lesser extent than specific binding). These non-specific interactions may be dependent on either similar amino acid sequences of ScKu1 and ScKu2 or the polyhistidine tag present on N-terminal regions. Importantly, cross-reactivity was significantly decreased when protein levels were reduced, indicating that cross-reactivity would be low at reduced protein levels, such as in cell extracts.

The most efficient immunodetection was observed with TB2 antibodies, however it was of interest to enhance specificity of antibody binding to ScKu proteins and to reduce background binding. To this effect, antibody dilutions were increased from 1:10000 to 1:20000, 1:50000 and 1:100000 as limited amounts of antibody tend to exhibit increased specificity for a particular target with highest affinity. Recombinant ScKu1 and ScKu2 were again used to determine immunodetection specificity of TB2 (data not shown) and final harvest bleed (FHB) antibodies (Figure 5.9).

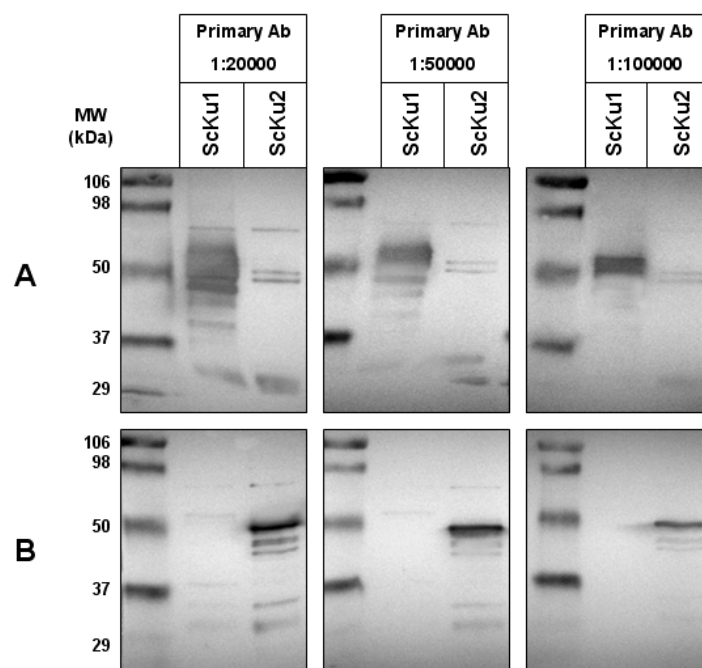


Figure 5.9 Immunodetection of ScKu1 and ScKu2 with varying dilutions of FHB polyclonal antibodies obtained from immunisation in rabbits. 5 pmol of recombinant ScKu1 and ScKu2 were subjected to SDS-PAGE and western blotting. Nitrocellulose membranes were then incubated in FHB polyclonal antibody at dilutions of 1:20000, 1:50000 and 1:100000 at 4 °C for 16 hours. Membranes were then exposed to anti-rabbit secondary antibody (1:10000) for 2 hours at 22 °C. **A.** Immunodetection with anti-ScKu1 antibodies. **B.** Immunodetection with anti-ScKu2 antibodies.

Increased dilutions of FHB anti-ScKu1 and anti-ScKu2 significantly increased specific binding to recombinant proteins and background detection of contaminating proteins was lowered. Although TB2 antibodies exhibited high-specificity, FHB antibodies yielded the most promising results and even at 1:100,000, binding of antibodies to specific proteins

remained high and cross-reactivity was very low. From this analysis it was decided that FHB antibodies would be most appropriate for future studies for immunodetection of endogenous ScKu1 and ScKu2 and immunoprecipitation assays as sensitivity and specificity was efficient even at elevated dilutions.

5.2.2 *S. coelicolor* and *E. coli* Cell Extract Analysis using ScKu and MtrA Antibodies.

The detection specificity of anti-ScKu1 and anti-ScKu2 had been proven using recombinant versions of *S. coelicolor* Ku proteins that had been used during the immunisation procedures. It was also of considerable interest to confirm that anti-ScKu1 and anti-ScKu2 could detect endogenous Ku proteins in *S. coelicolor* with relatively low cross-reactivity to other cellular proteins. Initial assays focused on the detection of ScKu1 and ScKu2 in various strains of *S. coelicolor* that had previously been produced (Handley, Bowater and Hutchings, unpublished data). Strain M145 was analysed alongside M145 deficient in ScKu1 (Δ *scku1*), M145 deficient in ScKu2 (Δ *scku2*) and M145 deficient in both ScKu1 and ScKu2 (Δ *scku1*/ Δ *scku2*). Following selected growth of strains in liquid culture (SFM) for 3 days at 30 °C, cells were sonicated and cell extracts were collected following centrifugation. Confirmation that cell extraction had been successful was achieved by Coomassie staining (Figure 5.10).

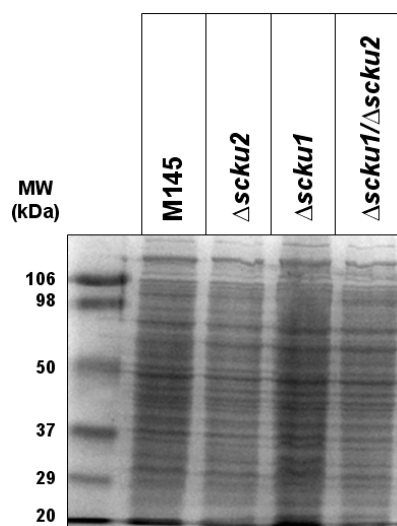


Figure 5.10 *S. coelicolor* cell extract analysis by Coomassie staining. *S. coelicolor* M145, Δ *scku1*, Δ *scku2* and Δ *scku1*/ Δ *scku2* cell extracts were harvested by sonication and centrifugation. Approximately 40 μ g of extract from each strain was subject to SDS-PAGE using 10% polyacrylamide gels and Coomassie staining to confirm the presence of proteins.

S. coelicolor cell extraction of proteins had been successful as indicated by the banding pattern visible in Figure 5.10. The observation that all strains exhibited a similar pattern of banding was reassuring as this implied all extracts were all derived from *S. coelicolor* and it

was likely all 4 were at the same stage of growth (stationary substrate mycelium stage). Subsequent analysis involved immunodetection of endogenous ScKu1 and ScKu2 in all 4 strains using appropriate conditions determined during recombinant protein immunodetection (Figure 5.10). As controls, antibodies from pre-immunisation bleeds were also analysed to confirm detection of proteins was determined by antibodies produced following ScKu immunisation. Also, an anti-MtrA antibody that had been produced by the same method as anti-ScKu antibodies was employed as a positive control. The MtrA protein is not implicated in NHEJ but is part of a two-component signal transduction system associated with signal recognition and adaptive responses (Zahrt and Deretic, 2000). The *S. coelicolor* operon SCO3011-SCO3013 has recently been identified as an orthologue of the *mtrAB-lpqB* operon of *M. tuberculosis* (Hutchings et al., 2006) with SCO3013 sharing 75% identity with the *M. tuberculosis* *mtrA* gene. The anti-MtrA antibody has been previously used to successfully detect MtrA in *S. coelicolor* M145 and, therefore, would be a suitable positive control for the strains analysed in this study.

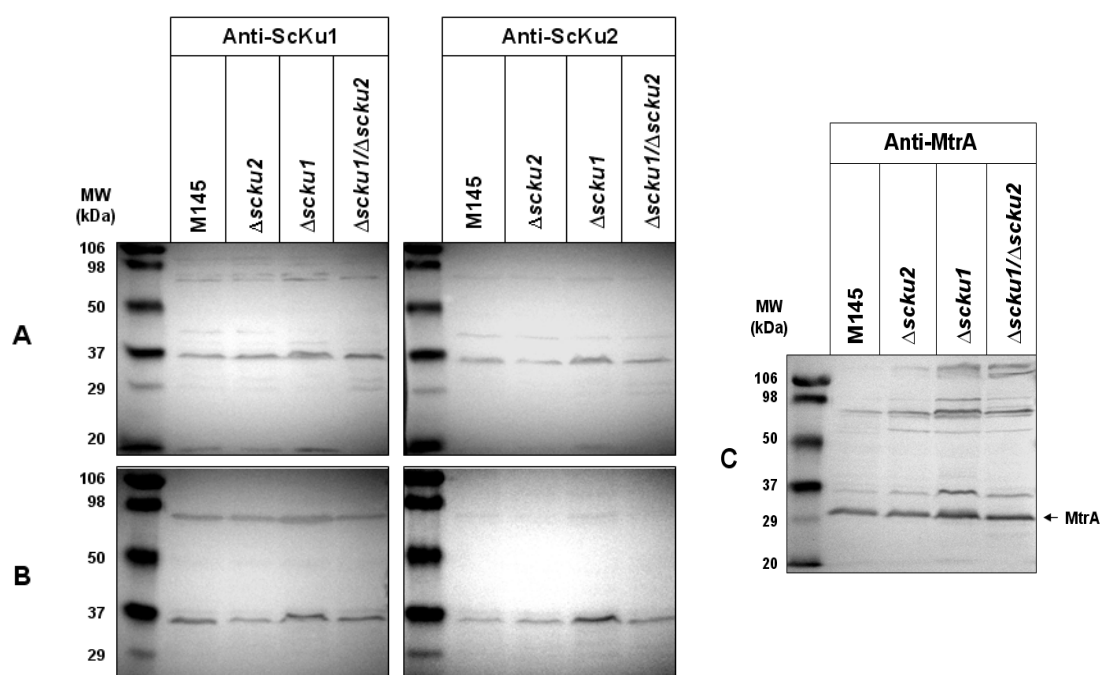


Figure 5.11 Endogenous *S. coelicolor* ScKu1, ScKu2 and MtrA immunodetection via polyclonal pre-immunisation and FHB antibodies. 50 μ g of *S. coelicolor* M145, $\Delta scku1$, $\Delta scku2$ and $\Delta scku1/\Delta scku2$ cell extracts were analysed by SDS-PAGE (10% polyacrylamide) and western blotting. **A.** Immunodetection with FHB anti-ScKu1 or anti-ScKu2 (1:100000). **B.** Immunodetection with pre-bleeds from rabbits immunised with ScKu1 or ScKu2. **C.** Immunodetection with FHB anti-MtrA.

Analysis of *S. coelicolor* strains with FHB anti-ScKu1 and anti-ScKu2 predominantly detected 2 protein species at approximately 80 and 35 kDa (Figure 5.11 A). As previous ScKu analysis had determined ScKu1 and ScKu2 have an electrophoretic mobility relative to

approximately 50 kDa it was expected that the observed bands were not ScKu. Analysis with pre-immunisation antibodies confirmed that the observed bands were not detected by antibodies from ScKu immunisation as both proteins were also detected during this analysis (Figure 5.11 B). Interestingly, the same bands were visible during MtrA immunodetection alongside bands representing MtrA (30 kDa). This analysis confirmed that native antibodies that were present in rabbits prior to protein immunisation or adjuvant addition were responsible for detecting the 2 protein species and that this occurrence was not rabbit specific as the same proteins were detected with antibodies from separate rabbits. As the anti-MtrA did successfully detect MtrA protein and no ScKu1 or ScKu2 were detectable in any *S. coelicolor* strains, it is likely that any ScKu present in these strains are poorly expressed under the growth conditions used and, therefore, are not visible following immunodetection.

ScKu specific antibodies had previously been used to visualise purified recombinant ScKu1 and ScKu2 proteins, however *S. coelicolor* cell extract analysis had failed to detect either endogenous versions in 4 separate strains. As it was considered appropriate to use ScKu antibodies to detect ScKu from cell extracts samples, it was decided to over-express ScKu in *E. coli* as implemented during protein purification processes (Section 2.4) and to use specific antibodies to analyse recombinant ScKu present in cell extracts (Figure 5.12).

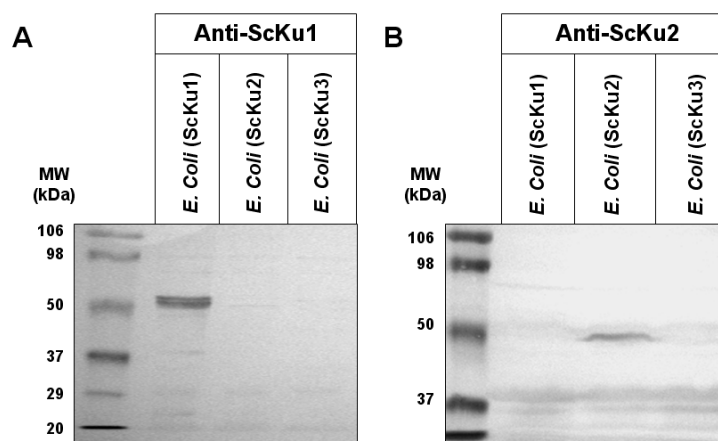


Figure 5.12 Immunodetection of recombinant ScKu1 and ScKu2 from *E. coli* cell extracts using polyclonal FHB antibodies. Recombinant ScKu1, ScKu2 and ScKu3 were over-expressed in *E. coli* BL21 pLysS strains from which cell extracts were harvested and analysed by SDS-PAGE (10% polyacrylamide), western blotting and immunodetection with appropriate antibodies. **A.** Immunodetection with FHB anti-ScKu1 on 5 µg of cell extract. **B.** Immunodetection with FHB anti-ScKu2 on 70 µg of cell extract.

Analysis of *E. coli* cell extracts illustrated that both anti-ScKu1 and anti-ScKu2 could be used successfully to detect ScKu proteins from cell extract material. Approximately 65 µg more of extract was used for anti-ScKu2 analysis as lower amounts generally gave low detection rates.

The probable reasons why less material was required during anti-ScKu1 immunodetection was that ScKu1 tends to express better during over-expression (Sections 4.2.1 and 4.2.2) and anti-ScKu1 appears to have higher sensitivity than anti-ScKu2 as shown earlier in this chapter. The successful detection of both ScKu1 and ScKu2 from *E. coli* cell extracts suggests that if native Ku proteins are present in *S. coelicolor* cell extracts (Figure 5.11) then they are present at low enough levels to avoid detection.

5.2.3 *S. coelicolor* Spore Extract Analysis using ScKu Polyclonal Antibodies.

S. coelicolor exhibit complex lifecycles that involve biochemically and morphologically differentiated cell types (Miguel et al., 2000). The *S. coelicolor* extracts investigated in Section 5.3.2 were taken from branching substrate mycelium thought to be at stationary phase and although Ku were not detected it is possible that expression increases at other lifecycle stages. Under stress, aerial filaments differentiate into spore-bearing structures that eventually produce hydrophobic dormant spore bodies (Hirsch and Ensign, 1976). As no cell division will occur in these dormant bodies and with no sister chromatid available, the repair of DSBs would not be able to proceed via homologous recombination and so NHEJ may predominate as the mechanism for repairing such breaks. To determine if ScKu1 or ScKu2 proteins are expressed in *S. coelicolor* spores, following growths on selective plates, spores were collected from the 4 *S. coelicolor* strains and extracts harvested using a FASTPrep technique to lyse cells. As with previous *S. coelicolor* extracts, 50 µg of extract was analysed by SDS-PAGE, western blotting and immunodetection with anti-ScKu1 and anti-ScKu2 (Figure 5.13).

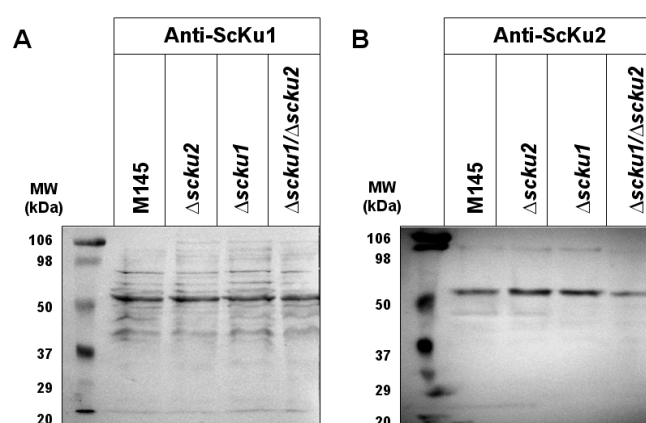


Figure 5.13 Immunodetection of native ScKu1 and ScKu2 from *S. coelicolor* spore extracts using polyclonal FHB antibodies. 50 µg of spore extracts from *S. coelicolor* M145, $\Delta scku1$, $\Delta scku2$ and $\Delta scku1/\Delta scku2$ were subjected to SDS-PAGE, western blotting and immunodetection with relevant primary antibodies at 1:100000 dilution. **A.** Immunodetection with FHB anti-ScKu1. **B.** Immunodetection with FHB anti-ScKu2.

Extracting spore extracts can be complicated due to the very durable nature of spores. However, Coomassie analysis (data not shown) confirmed the FASTPrep technique had successfully broken up spore cells and extracts rich in protein had been attained. Similarly to *S. coelicolor* cell extract analysis, spore analysis identified very little difference in proteins present between the strains studied regardless of choice of antibody during immunodetection. As seen in previous *S. coelicolor* cell extract analysis (Figure 5.11), a single-band was consistently detected in all samples with both antibodies, however this time the unknown protein was larger at approximately 55 kDa. The observation of this protein species, and loss of the 2 bands clearly identified previously indicates protein expression levels are different in *S. coelicolor* spores compared to stationary substrate mycelium. Importantly, the absence of ScKu proteins indicates levels are too low for detection, suggesting that either these proteins can function at very low concentrations, or that they are not required at this life stage. It is also possible expression may be induced following DSB formation and that levels will only be detectable in damage-induced cells.

5.3 Resolving ScKu Protein Macromolecular Interactions.

The 3 ScKu proteins had all exhibited DNA binding activity that would be fundamental to their hypothesised role during NHEJ in *S. coelicolor*. Importantly, ScKu displayed this activity independent of each other, and considering all Ku are believed to bind to DNA as dimeric complexes, it would be assumed that these proteins conform to the currently accepted model of prokaryotic Ku (Pitcher et al., 2007) and form homodimers. Whereas a high proportion of bacteria possess single copies of *ku* genes (Rocha et al., 2005; Wilson et al., 2003), some bacteria, including *S. coelicolor* contain multiple *ku* genes (Kobayashi et al., 2008), providing the possibility that these Ku proteins may function as heterodimers in a similar fashion to eukaryotic Ku. Due to the structural similarity observed by individual Ku monomers, it is not unreasonable to suppose that different Ku from the same species could also interact with each other. To understand this possibility further, biochemical analysis of the 3 purified recombinant ScKu proteins by dynamic light scattering was carried out to ascertain if multiple species of protein were present in preparations. Co-immunoprecipitation assays were also developed using antibodies described in Section 5.2 that provided the opportunity to assess the ability for ScKu to interact and also to discover if ScKu proteins could interact with the DNA ligase implicated in *S. coelicolor* NHEJ (ScLigD). As a final approach to determine protein-protein interactions using precipitation assays, the SMB DNA binding assay was adapted to resolve if proteins that had previously been found to interact poorly with specific DNAs could interact more efficiently if a protein found to bind well to the DNA was present in the reaction.

5.3.1 Dynamic Light Scattering Analysis of ScKu.

Dynamic light scattering (DLS) is a technique that can be used to determine the size distribution of small particles in solution. When a monochromatic and coherent beam of light such as a laser is scattered by particles, a time-dependent fluctuation in the scattering intensity can be observed due to the movement of particles by Brownian motion. This movement causes the distance between scatters in the solution to change with time and these changes can be quantified to produce information regarding the movement of the particles.

Polydispersity refers to the level of homogeneity of the sizes of particles in a solution. High homogeneity (for example when percent polydispersity is less than 15%) means that particles can be considered to be virtually identical, however when percent polydispersity rises above 30%, the particle population can be considered to contain species of different sizes. When particle species are present that are more than a factor of 2 different in size, additional peaks are resolved during DLS indicating that the sample is heterogeneous and that distinct populations are present. Heterogeneity can also be caused by the presence of different species that are unable to be resolved by DLS, meaning that species are present with sizes less than a factor of 2 relative to other species in the same solution. In such instances, a single peak will arise that exhibits higher polydispersity (indicated by a wide peak) compared to a solution that only contains a single particle species (indicated by a narrow peak). The data shown in Table 5.2 contains DLS data following analysis of ScKu1, ScKu2 and ScKu3. The ‘% polydispersity’ values were calculated from the widths of the main peaks (Table 5.2).

Table 5.2 DLS data obtained from analysis of ScKu proteins. 12 µl of 1 mg / ml protein in 1 x protein storage buffer (Table 2.11) were measured at an acquisition time of 10 seconds for a total of 20 acquisitions. Values representing ‘% polydispersity’ were calculated from the main peak in each measurement. Values < 15% indicate the presence of 1 species, values between 15 – 30% indicate monomers with increasing amounts of dimer present and values > 30 % indicate highly heterogeneous species. M = monomer, D = dimer, T = trimer.

Protein	‘% Polydispersity’	Protein species
ScKu1	49.2	M/D/T
ScKu2	21.7	M/D
ScKu3	18.9	M/D

The percent polydispersity values acquired from DLS measurements suggest that all 3 ScKu proteins are present as more than a single species. For both ScKu2 and ScKu3, the width of the main peak produced values between 15 – 30% polydispersity, which indicates both monomer and dimer species exist under the conditions stated (Section 2.12). Interestingly, ScKu1 displayed higher percent polydispersity that is indicative of a protein existing in

multiple forms such as monomers, dimers and trimers. As no other Ku proteins have been documented as existing in a trimeric form, it is likely this value does not truly represent the ScKu1 species present. Overall, the data implies that the protein preparations are likely to contain both monomer and dimer molecules.

5.3.2 Co-Immunoprecipitation of *S. coelicolor* NHEJ Components.

For characterising *in vitro* protein-protein interactions of *S. coelicolor* Ku and other NHEJ components, a co-immunoprecipitation (Co-IP) assay was developed that employed ScKu specific antibodies (Section 5.2) to precipitate target proteins physically associated with the antibody specific ScKu bait. Preliminary assays were performed to optimise the Co-IP assay (Section 2.9) and to reduce non-specific protein-bead and protein-antibody interactions from occurring. Assays then focused on comparing precipitation of target proteins with or without the inclusion of ScKu bait proteins to which the antibody present in reactions had been raised against. It was assumed that increases in precipitation of target proteins when the ScKu bait protein was present would be a direct consequence of a protein-protein interaction between ScKu bait and target protein. Co-IP assays were primarily performed to quantify if different ScKu proteins could form stable heterodimer complexes and also to resolve if either ScKu1 or ScKu2 could enhance precipitation of ScLigD (Figure 5.14).

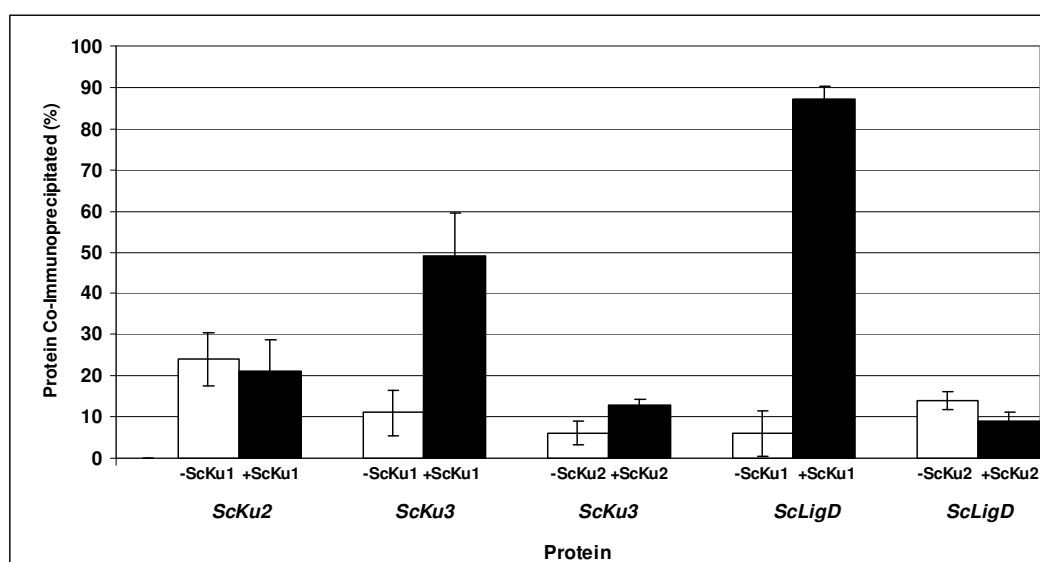


Figure 5.14 *In vitro* Co-IP results investigating *S. coelicolor* NHEJ component protein-protein interactions. Assays were performed to determine target protein Co-IP with or without antibody specific bait protein (normal font). Values representing percentage target protein (in italics) precipitated in both instances were calculated using 'ImageJ'. Bars in white indicate precipitated target protein without ScKu bait protein (non-specific Co-IP) and bars in black indicate precipitated target protein with ScKu bait included (specific Co-IP). Averages were taken from a minimum of 3 experiments. Standard errors are shown.

Table 5.3 Statistical analysis of results obtained during Co-IP assays investigating protein-protein interactions between *S. coelicolor* NHEJ components. Two-tailed t-test analysis was performed on data obtained from assays shown in Figure 5.14. The percentages of protein co-immunoprecipitated with or without incubation with antibody-specific protein were calculated and then directly compared to resolve differences in values.

Protein associated with PGB	Target protein	<i>t</i> value	Degrees of freedom (of <i>t</i>)	Critical values of <i>t</i> (<i>p</i> = 0.05)	Outcome
ScKu1	ScKu2	0.3	4	2.78	Not significant
ScKu1	ScKu3	3.2	4	2.78	Significant
ScKu2	ScKu3	2.7	4	2.78	Not significant
ScKu1	ScLigD	12.5	2	4.30	Significant
ScKu2	ScLigD	1.8	2	4.30	Not significant

The data from Co-IP assays indicated that ScKu1 was able to precipitate both ScKu3 and ScLigD, however any interaction with ScKu2 was not apparent. ScKu2 was also shown not to significantly increase precipitation of ScKu3 or ScLigD, implying that the interactions shown by ScKu1 is specific for this protein only. The considerable co-precipitation of ScLigD by ScKu1 is an interesting result. Although in the context of prokaryotic NHEJ, an interaction between the DSB detecting ScKu and the DNA ligating ScLigD would seem plausible, studies of *M. tuberculosis* NHEJ proteins suggested that MtKu interacts with the polymerase domain of MtLigD during recruitment of the ligase (Pitcher et al., 2007). As ScLigD does not contain a polymerase domain, it may be hypothesised that for a system similar to NHEJ in *M. tuberculosis*, ScKu would interact with *S. coelicolor* proteins that show high homology to the MtLigD polymerase domain (ScPrim1, ScPrim2 and ScPrim3) (Sayer, 2008). Though interactions to these proteins are yet to be elucidated, the observed interaction between ScKu1 and ScLigD indicates differences in interactions between *S. coelicolor* and *M. tuberculosis* NHEJ proteins exist.

The observation that between the ScKu proteins only ScKu1 and ScKu3 appear to interact is intriguing. It may be expected due to the genomic locations of these proteins that chromosomal ScKu1 and ScKu2 would be most likely to interact compared to the isolated ScKu3 on plasmid SCP1. For conformation of ScKu1/ScKu2 interactions, ideally the roles of these 2 proteins would also have been reversed with ScKu2 as bait and ScKu1 as target protein. Unfortunately, due to the appearance of a band at the migratory position of ScKu1 during electrophoresis (assay antibody background (Appendix IV)) the quantification of ScKu1 was not possible. *In silico* sequence analysis in 3 other *Streptomyces* species (Appendix V) has revealed that in these other species, 2 contain 2 *ku* genes (showing highest homology to ScKu1 and ScKu3) and 1 species contains only 1 *ku* gene with highest homology to ScKu1. This information strongly suggests ScKu1 is capable of functioning

alone as a homodimer, but also that potentially it may interact with ScKu3 in 2 out of the 3 alternative species analysed. The absence of ScKu2 in any of the 3 other species implies *S. coelicolor* has at some point acquired this gene and that, although functionally similar to ScKu1 and ScKu3, it is unlikely to form complexes with these proteins and is not essential for all *Streptomyces* species. Hence, Co-IP and *in silico* data indicate that prokaryotic Ku may not be restricted to forming homodimers as currently thought. Although it is apparent all 3 ScKu can bind to DNA individually, and therefore supposedly as homodimers, the observed interaction between ScKu1 and ScKu3 during this study, and the presence of these 2 *ku* genes in 2 other *Streptomyces* species indicates the ability for these Ku to heterodimerise. The ability of ScKu1 and ScKu3 to form both homodimer and heterodimer complexes may allow these proteins to participate in different cellular activities, either during repair of different types of DSB, or alternative DNA metabolism processes.

5.3.3 *In vitro* SMB DNA and Protein Interaction Analysis.

The use of antibodies specific for ‘bait’ proteins during Co-IP is an effective method for determining if 2 different proteins can interact. However, potential drawbacks exist regarding the disruptive influence that a relatively large immunoglobulin molecule may have upon protein function. This effect may be exacerbated when polyclonal antibodies that cannot be directed against non-obtrusive epitopes are used. Therefore, the ability to support data collected from such Co-IP analysis is highly advantageous. The differential DNA binding activity displayed by *S. coelicolor* NHEJ components during SMB assays (Figure 4.4) provided an opportunity to develop a more natural Co-IP assay that would utilise the proteins DNA binding activity to allow precipitation, rather than using protein-specific antibodies. Effectively, a protein that had been shown to bind poorly to a specific DNA molecule was incubated with a protein that had been shown to bind efficiently to the same DNA. If protein precipitation increased in the presence of the efficient DNA binding protein, it was thought an interaction was occurring between both proteins. Initial SMB DNA and protein interaction assays focused on possible interactions between the 3 Ku proteins of *S. coelicolor* (Figure 5.15).

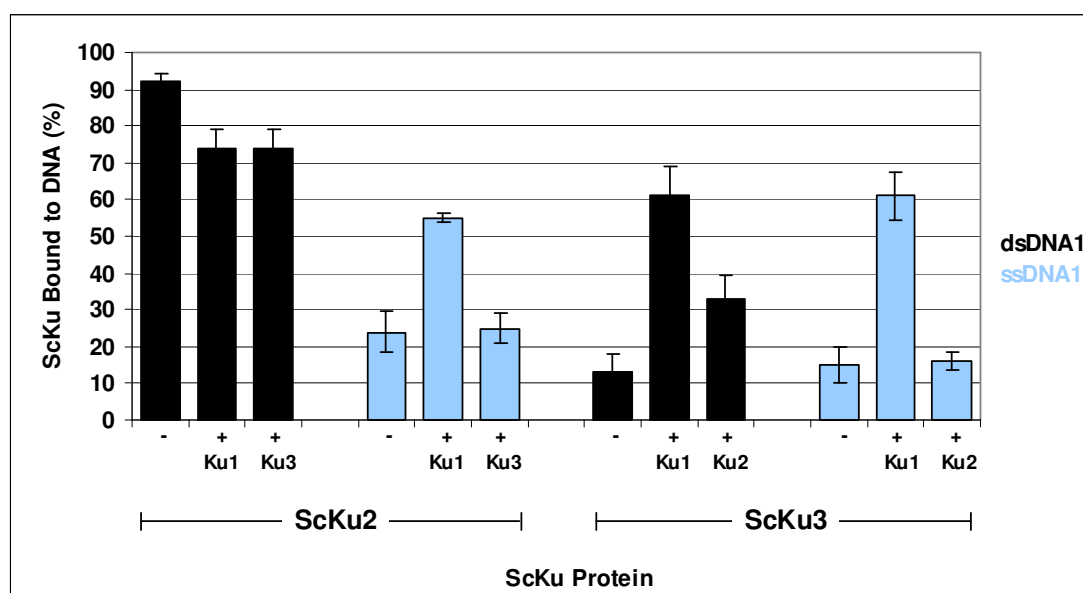


Figure 5.15 *In vitro* SMB DNA binding assay results investigating protein-protein interactions between ScKu proteins. Different ScKu were incubated prior to exposure to dsDNA1 (black) or ssDNA1 (blue) (Table 2.7) using assay conditions described in Section 4.3.1. Band intensities were measured using ‘ImageJ’ and percentage binding calculated. Results for ScKu2 and ScKu3 are shown without (-) or including stated ScKu proteins. Averages were taken from at least 4 experiments. Standard errors are shown.

Table 5.4 Statistical analysis of results obtained during SMB DNA and protein assays investigating ScKu protein interactions. Two-tailed t-test analysis was performed on data obtained from assays shown in Figure 5.15. The percentages of ScKu2 and ScKu3 precipitated (with dsDNA or ssDNA) with prior incubation of alternate ScKu protein were compared with percentage precipitated without prior incubation.

Protein combination	DNA	<i>t</i> value	Degrees of freedom (of <i>t</i>)	Critical values of <i>t</i> (<i>p</i> = 0.05)	Binding difference
ScKu2 (ScKu1)	dsDNA	4.0	16	2.12	Significant
ScKu2 (ScKu3)	dsDNA	3.9	16	2.12	Significant
ScKu2 (ScKu1)	ssDNA	3.3	13	2.16	Significant
ScKu2 (ScKu3)	ssDNA	0.03	14	2.15	Not significant
ScKu3 (ScKu1)	dsDNA	5.4	9	2.26	Significant
ScKu3 (ScKu2)	dsDNA	2.4	9	2.26	Significant
ScKu3 (ScKu1)	ssDNA	5.7	10	2.23	Significant
ScKu3 (ScKu2)	ssDNA	0.2	9	2.26	Not significant

The results shown in Figure 5.15 and subsequent statistical analysis (Table 5.4) demonstrate that the presence of alternate ScKu proteins did affect precipitation depending on which other ScKu protein was present. Previous DNA binding analysis had identified ScKu2 bound well to dsDNA, but comparatively less to ssDNA. When either ScKu1 or ScKu3 was present, ScKu2 precipitation with dsDNA was shown to decrease. This may be explained by competition between different ScKu proteins that would reduce the amount of DNA available for ScKu2 to bind to. Thus, this would indicate that ScKu2 does not interact with ScKu1 or

ScKu3 as these proteins appear to reduce access of ScKu2 to the dsDNA. The results attained representing ScKu2 precipitation when ssDNA was used as substrate appear to contradict this observation as although no difference was observed when ScKu3 was present (as would be expected as neither ScKu2 nor ScKu3 bind efficiently to ssDNA), the inclusion of ScKu1 significantly increased ScKu2 precipitation. This would imply an interaction between ScKu1 and ScKu2 is possible, although as precipitation with dsDNA and previous Co-IP assays (Section 5.3.2) did not detect such interaction, other factors may be responsible for this observation. The results concerning precipitation of ScKu3 were very encouraging as there was high agreement with previous antibody-dependent Co-IP analysis. ScKu1 appeared to strongly increase precipitation of ScKu3 with both dsDNA and ssDNA, again suggesting an interaction between these 2 proteins. Although the inclusion of ScKu2 did increase precipitation of ScKu3 with dsDNA, this increase was not as prominent as that observed when ScKu1 was present.

In vitro SMB DNA binding assays had also revealed that ScLigD, an ATP-dependent DNA ligase from *S. coelicolor* believed to function during NHEJ (Aravind and Koonin, 2001; Doherty et al., 2001), exhibited contrasting DNA binding activity to ScKu and MtKu proteins. As ScLigD bound with relatively low affinity to dsDNA compared to ScKu1, ScKu2 and MtKu, an opportunity existed to develop the SMB assay to allow identification of protein-protein interactions with Ku proteins, an interaction that may be crucial during NHEJ. For these assays, ScKu were incubated with dsDNA and unbound protein was removed in wash steps, followed by the introduction of ScLigD to each individual reaction. Reactions would then be incubated and bound protein precipitated as usual, however, the presence of ScLigD would be dependent on interactions with the Ku protein attached to the DNA and not direct interactions with the dsDNA. Assays were carried out that investigated whether ScLigD would be precipitated following incubation of ScKu1, ScKu2 or MtKu with dsDNA.

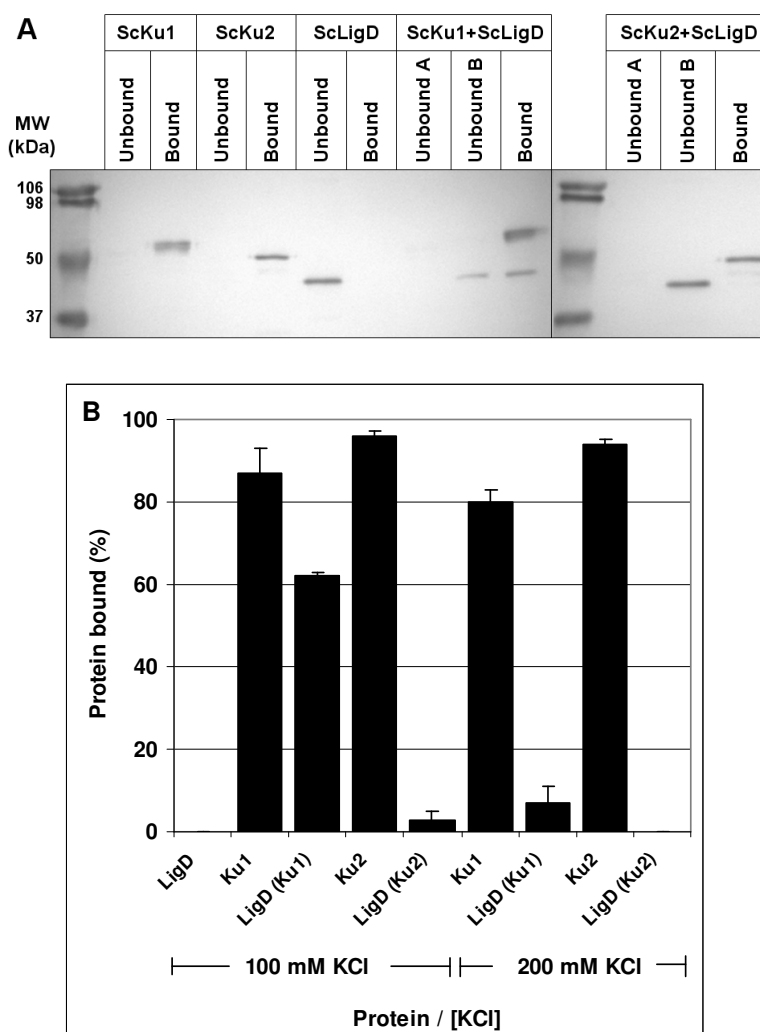


Figure 5.16 *In vitro* SMB DNA binding assay results investigating protein-protein interactions between prokaryotic Ku and ScLigD. Prokaryotic Ku were initially incubated with dsDNA1 (Table 2.7) followed by addition of ScLigD that has low affinity for dsDNA. Precipitation of ScLigD would be dependent on interactions with ScKu. Averages were taken from a minimum of 3 experiments. Standard errors are shown. **A.** Example of SMB results (performed at 100 mM KCl) following SDS-PAGE, western blotting and immunodetection. Proteins can be distinguished by differences in electrophoretic migration. **B.** Summary of protein-protein interaction assays performed at 100 and 200 mM KCl. Proteins in brackets were incubated with DNA prior to ScLigD addition.

The results obtained from Ku-ScLigD interaction studies using the modified SMB DNA binding assay (Figure 5.16) indicated that when ScKu1 was bound to dsDNA, precipitation of ScLigD was significantly higher than when no Ku was present ($t = 62$ ($p < 0.05$)) or when ScKu2 was bound to dsDNA ($t = 12$ ($p < 0.05$)) (at 100 mM KCl). These findings imply that ScKu1 can facilitate ScLigD association with dsDNA, probably via protein-protein interactions that ScKu2 is less capable of performing. The difference in strength between ScKu1-ScLigD and ScKu2-ScLigD interactions is highlighted as when ionic strength was increased to 200 mM KCl, no ScLigD was precipitated in the presence of ScKu2, however

ScLigD was still precipitating when ScKu1 was present. Assays were also performed to investigate potential interactions between MtKu and ScLigD, however no interactions were detected (data not shown).

This data strongly agrees with other Co-IP data obtained during this study as ScKu1 appears to interact directly with ScLigD to facilitate association with DNA. If ScLigD is indeed responsible for ligation of DSBs during NHEJ in *S. coelicolor* as expected, the findings from previous assays that this ligase binds poorly to dsDNA would imply that additional components are necessary in order for ligation to proceed. The introduction of ScKu1 dramatically increased the amount of ScLigD attached to dsDNA and this apparent interaction may reflect the specific *in vivo* function of ScKu1 in enhancing the opportunity for DSB ligation. Why the addition of ScKu2 resulted in less considerable increases in ScLigD precipitation is not obvious. Perhaps additional components are required that enhance ScKu2-ScLigD interactions or, alternatively, ScKu2 may facilitate ligation by a different ligase during NHEJ.

5.4 Discussion.

Two approaches were taken that will ultimately allow for detailed *in vivo* analysis of ScKu1 and ScKu2 activity in *S. coelicolor*. The first approach was to integrate polyhistidine tagged recombinant versions of *sku1* and *sku2* using phage-derived vectors pMS82 and pSET152 that integrate directly into separate locations on the *S. coelicolor* chromosome. Primers were designed that would amplify not only the appropriate *sku1* and *sku2* genes, but also approximately 200 bp of upstream DNA that would contain promoter regions to allow for natural levels of protein expression. Prior to gene amplification, restriction digest analysis of pMS82 and pSET152 was performed and electrophoretic analysis of the resulting fragments agreed with the theoretical fragment molecular sizes. PCR amplification of both *sku1* and *sku2* from *S. coelicolor* chromosomal DNA proved problematic and cosmid DNA was chosen as an alternative template. Once appropriate reaction conditions had been attained, *sku2* amplification was successful, however PCRs specific for *sku1* failed to produce any product. Although several possibilities exist for why *sku1* amplification was unsuccessful, the high G/C content of *S. coelicolor* may have impaired primer function and promoted events such as self-folding. As no DNA had been amplified regarding *sku1*, all further cloning concentrated exclusively on *sku2*. PCR products amplified by the proofreading Pfu polymerase were gel-extracted, purified and subsequently cloned into pCR®-Blunt II-TOPO vectors prior to eventual sub-cloning into pMS82 and pSET152. Analysis by colony PCR and restriction digestion confirmed the desired gene products had been inserted appropriately and

confirmation was attained via sequencing of pCR[®]-Blunt II-TOPO vectors containing *sku2* inserts. Integration vectors and pCR[®]-Blunt II-TOPO vectors containing inserts were digested with appropriate restriction enzymes and, following gel extraction of relevant DNAs, *sku2* inserts were ligated into pSET152 and successful cloning was confirmed via colony PCR and restriction digests and sequencing.

The production of antibodies specific for individual ScKu provides the opportunity to investigate several new aspects of the functions of these proteins. Previous immunodetections had utilised the polyhistidine tags of recombinant ScKu and so the detection of native ScKu in *S. coelicolor* and the selective precipitation of specific recombinant ScKu *in vitro* was not possible with anti-his antibodies. Antibodies specific for ScKu1 (anti-ScKu1) and ScKu2 (anti-ScKu2) were developed by immunisation of rabbits with purified recombinant proteins from which polyclonal antibodies were eventually harvested.

Preliminary analysis of antibodies was performed by probing known amounts of recombinant ScKu1 and ScKu2 with pre-immunisation bleeds, a primary test bleed (TB1) taken following the first immunisation and a second test bleed (TB2) that was taken after 2 immunisation events. All antibodies were initially used at a dilution of 1: 10000 and varied amounts of ScKu protein were analysed. Although initial results were discouraging due to a high degree of non-specific detection of contaminating proteins, increased antibody dilutions significantly reduced cross-reactivity. It was determined that 1:100,000 dilutions, both antibodies provided the relatively clean results and could efficiently detect 5 pmol of recombinant protein. As anticipated, final harvest bleed (FHB) antibodies exhibited the most efficient detection of ScKu proteins as these samples would contain the highest levels of ScKu-specific antibody molecules. FHB antibodies were then employed in the detection of native ScKu1 and ScKu2 from *S. coelicolor* cell extracts. *S. coelicolor* strains available for analysis were: M145, M145 Δ ScKu1, M145 Δ ScKu2 and M145 Δ ScKu1/ Δ ScKu2. Cell extracts obtained from these strains grown in liquid culture for 3 days were analysed by Coomassie staining to confirm harvesting had been successful and, once confirmed, extracts were subject to SDS-PAGE, western blotting and immunodetection with pre-bleed antibodies, FHB anti-ScKu1, anti-ScKu2 and FHB anti-MtrA (used as a positive control). Two protein species at approximately 80 and 35 kDa were consistently detected with all 3 protein antibodies including ScKu pre-bleeds, indicating antibodies native to the immunised rabbits were present that were specific for *S. coelicolor* proteins of these sizes. Although anti-MtrA did detect an additional band at 30 kDa thought to be MtrA, neither ScKu FHB antibodies revealed additional bands at the expected molecular weights of the ScKu proteins. Therefore, it was assumed any ScKu1 or

ScKu2 present from the analysed cell extracts was at an amount too low for immunodetection. To confirm the ability of FHB anti-ScKu1 and anti-ScKu2 to detect relevant proteins from cell extracts, *E. coli* over-expressing both recombinant proteins were examined and both proteins were detected successfully. As *S. coelicolor* exist in multiple biochemically different cell types, it was postulated that expression of ScKu would vary depending on the exact stage of this cycle. As spores are dormant bodies that do not exhibit cell division, it was believed ScKu expression may be increased as NHEJ may be more prevalent in these non-dividing cell types. Spore extractions were obtained and again subjected to SDS-PAGE, western blotting and immunodetection with FHB anti-ScKu1 and ScKu2. A constant single-band was detected in all samples including anti-MtrA but no other bands appeared to show variation between the different strains. Interestingly, the band that was observed exhibited different electrophoretic mobility to the 2 bands seen in stationary substrate mycelium extracts. This indicates differences in protein expression between the 2 life cycles studied, however there was still no evidence of expression of ScKu1 or ScKu2. Potentially, expression of both ScKu1 and ScKu2 are dependent upon the formation of DSBs, and therefore expression may only be detectable in damage-induced cells, a possibility that requires further investigation.

Anti-ScKu1 and anti-ScKu2 immunoglobulins were also used during *in vitro* co-immunoprecipitation (Co-IP) analysis investigating protein-protein interactions between ScKu proteins and also ScLigD. Previous DNA binding analysis had identified all 3 ScKu were able to bind to DNA independently, a factor that strongly implied these proteins were able to form homodimer complexes. Additional DLS data provided further evidence of this ability as multiple species of ScKu were detectable in protein preparations as illustrated by the percent polydispersity of peaks representing ScKu proteins. Even though these proteins may function as homodimers, it was proposed that due to the high structural similarity of Ku proteins, Ku from the same organism may also heterodimerise, as evident with eukaryotic Ku70/80. Co-IP analysis provided strong evidence that to a degree, ScKu do elicit the ability to interact with each other. Although ScKu2 failed to precipitate ScKu1, ScKu3 or ScLigD, an interaction was apparent between ScKu1 and ScKu3. ScKu1 was also shown to interact with ScLigD, a finding which was not anticipated considering the characterised interaction between MtKu and the polymerase domain of MtLigD (Pitcher et al., 2007). Another surprising feature of this interaction was that it appeared to be of considerable strength and indeed, ScKu1 was consistently able to precipitate more ScLigD than ScKu3. It would be thought that the formation of a specific heterodimer complex (as may be supposed between ScKu1 and ScKu3) would be stronger than an interaction between 2 unrelated proteins. An important factor that may have significantly reduced ScKu1 and ScKu3 interactions would be that a high proportion of these proteins may already have existed as homodimers prior to

immunoprecipitation assays, which would inevitably negatively affect heterodimer formation. Perhaps to appreciate a true comparison between ScKu1-mediated precipitation of ScKu3 and ScLigD, ensuring both Ku preparations were in a monomeric state prior to Co-IP would be important.

To further quantify protein-protein interactions between *S. coelicolor* NHEJ components, a modification was made to the existing SMB DNA binding assay whereby precipitation of proteins would either be dependent on their ability to bind to a specific DNA molecule or to interact with a protein that can bind to the DNA. Although ScKu1 had been shown to bind to both dsDNA and ssDNA, ScKu2 was recognised as having low affinity towards ssDNA, ScLigD could not bind to dsDNA and ScKu3 bound poorly to both DNA substrates (Section 4.3.1). Therefore, if binding of these proteins to the stated DNAs was increased in the presence of other proteins, it would be assumed that precipitation was dependent on protein-protein interactions. Similarly to previous antibody-dependent Co-IP results, the presence of ScKu1 significantly increased ScKu3 precipitation with both DNAs. Results also implied that ScKu2 could not interact with either ScKu1 or ScKu3 as precipitation of ScKu2 when dsDNA was available was reduced when these proteins were present. It was assumed this was an example of competing proteins reducing available DNA substrates by effectively blocking ScKu2 binding. Two results appeared to conflict with the concept that ScKu2 did not interact with other ScKu; the fact that precipitation of ScKu2 when ssDNA was substrate was increased when ScKu1 was present, and the increased precipitation of ScKu3 to dsDNA when ScKu2 was present (albeit considerably less than when ScKu1 was present). It is possible that ScKu2 can indeed interact with ScKu1 and ScKu3, and that this interaction is weaker than ScKu1-ScKu3. However, these observations may also be determined by alterations to the DNA following protein binding that allow a protein that has previously shown poor affinity to that particular DNA to bind as a consequence. *In silico* data (Appendix V) reinforces the implication that ScKu1 and ScKu3 may function together as 2 out of 3 other *Streptomyces* species that have been sequenced contain homologues for these 2 proteins, but none contain ScKu2 homologues. Overall, the data indicates that as well as functioning as homodimers, ScKu1 and ScKu3 are likely to form functional heterodimers. The ability for ScKu2 to interact with other Ku proteins remains less convincing and it is possible that this protein may either act independently as a homodimer or it may be that this protein does not even contribute to NHEJ in *S. coelicolor*.

Finally, the SMB DNA and protein interaction assay provided an opportunity to detect if prokaryotic Ku could interact with ScLigD. Observations on DNA-binding interactions towards dsDNA and ssDNA had shown ScLigD binding to dsDNA was very low, however,

prokaryotic Ku bound with high affinity towards this substrate. Assays that consisted of a pre-Ku incubation with dsDNA, followed by the addition of ScLigD were performed and it was assumed that if ScLigD was precipitated, then Ku were directly responsible for this occurrence. It was found that although precipitated ScLigD was very low for ScKu2, and non-existent for MtKu, the presence of ScKu1 significantly increased ScLigD association with DNA. This apparent protein-protein interaction was almost abolished when ionic strength was increased, indicating that electrostatic forces are an important factor. A separate view on this observation may be that binding of ScKu1 induced conformational changes in the dsDNA to which ScLigD had increased affinity towards and so additional investigations regarding this interaction are required. Why ScKu2 appeared to interact with ScLigD much less efficiently remains unclear but may reinforce the possibility ScKu2 is not important for NHEJ in *S. coelicolor*. Alternatively, additional components may be required for a stable interaction to occur, or ScKu2 may even recruit an alternative ligase to DSBs for repair to proceed.

CHAPTER 6

General Discussion

The interactions that occur between proteins and nucleic acid molecules are intricate and dependent upon several factors. As experimental techniques continue to advance, the general understanding of the precise aspects that govern different modes of protein-DNA interactions have become clearer. A significant element of this study has been to develop and optimise alternative *in vitro* approaches for investigating DNA binding protein activity. It was anticipated that these approaches could be used to deconvolute currently published data regarding *in vitro* DNA binding activity of the intensely studied human p53 protein and also to investigate functions of the comparatively unknown Ku proteins from *S. coelicolor*. Although non-related, both p53 and Ku are heavily implicated in the regulation of DNA repair and the ability to bind directly to DNA encompasses a crucial activity of both proteins.

6.1 Characterising Human p53-DNA Interactions.

The preservation of genomic integrity is a fundamental activity performed by p53 in a diverse range of multi-cellular organisms. Although the heavily documented transactivation function of p53 is integral to genomic maintenance, p53 also exerts its effects independent of transcriptional regulation of target genes. The proposition that p53 may behave as part of a global surveillance system that is responsible for the detection of unstable genomic regions further increases the apparent importance this protein has on preventing potentially catastrophic occurrences (Walter et al., 2005). Direct interactions with DNA are crucial for p53 activity, as illustrated by the often severe consequences following mutation events in core DNA-binding regions (Shu et al., 2006). In common with many DNA binding proteins, p53 is capable of multiple modes of DNA interactions that show considerable flexibility. The ability to recognise and bind to consensus DNA sequences is instrumental in the orchestration of gene expression of specific genes that regulate cell cycle arrest, DNA repair and apoptosis. This interaction is determined by both sequence and topology of the response element and is heavily dependent on the core DNA-binding domain (Gohler et al., 2002; Kim and Deppert, 2006; Weinberg et al., 2005). In addition, p53 can also interact non-specifically with linear DNA, a feature postulated to enhance the location of specific DNA regions by facilitating

localisation to DNA (McKinney et al., 2004), and also sequence-independently to certain non-B DNA structures (Degtyareva et al., 2001; Stansel et al., 2002). The structure-dependent mode of binding may allow p53 to partake in processes such as DNA repair independent of transcription by recruiting repair components to dangerous chromosomal sites (Reed et al., 1995). The reported ability of p53 to bind to structures formed by CTG•CAG tracts with affinity equal to consensus DNA binding is of considerable interest (Walter et al., 2005). Similar structures are thought to arise in intrinsically unstable TRS tracts that are highly prone to structural rearrangements (Kulkarni and Wilson, 2008; Rass et al., 2007), a feature believed to be the underlying cause of several human hereditary neurological diseases. It was proposed from these findings that latent p53 in unstressed cells may be responsible for maintaining genomic integrity by regulating unstable genomic regions.

This study aimed to use an alternative *in vitro* DNA binding assay to analyse p53-DNA binding to various DNAs to allow detection of sequence-specific, structure-specific and non-specific modes of DNA interaction. Ultimately, a priority was to determine if the currently published data regarding p53 binding to hairpins comprising of CTG•CAG could be replicated using a different assay system. The deployment of a new *in vitro* approach to study p53-DNA binding would be highly beneficial considering the restraints of more traditional methods. *In vivo* analysis of p53-DNA binding was also investigated during this study that concentrated on distinguishing between sequence-specific and non-sequence-specific binding.

Prior to *in vitro* protein-DNA binding analysis the analysis of assay components was crucial to ensure consistent and accurate results were attainable. For SMB assays, the production of stable oligonucleotides that were homogenous in composition was requisite and methodical optimisations to the annealing procedure meant the appropriate DNAs were eventually attainable. Both *in vitro* DNA binding assays described in this study employed magnetic beads that were used to precipitate protein-DNA complexes. Factors including bead capacity and the propensity to form non-specific interactions with protein or DNA elements required characterisation before experimentation. Analysis of SMB DNA capacity indicated that the values specified by the manufacturer were relatively accurate although generally DNA attachment was slightly less than stated. Importantly, it was recognised that, to achieve these values, DNA must be added in excess (by approximately 3 times). Once attached, the tight biotin-streptavidin interaction ensured that DNA remained associated with SMBs even following several wash steps. An important feature of SMBs that was quickly apparent was the non-specific interactions that occurred with some proteins. This interaction was protein dependent and the magnitude of protein precipitated when no DNA was present varied between proteins. The introduction of a free-biotin incubation, addition of BSA to assay

buffers and increased KCl in reactions considerably reduced non-specific interactions to less than 20%, which, although not ideal, allowed differences in p53-DNA binding to be resolved. In contrast, analysis of PGBs demonstrated p53 precipitation only occurred in the presence of p53 specific antibodies under standard assay conditions. Capacity analysis of PGBs was also found to agree with the recommended protocol. Although the PGB assay did not require modification, SMB analysis proved to be essential and reinforces the necessity to scrutinise assay components before commencing experiments of this nature.

The *in vitro* SMB DNA binding assay was used to successfully detect 3 distinct modes of p53-DNA interactions. Replicating cellular conditions *in vitro* is a complicated process that is far from being resolved, however mimicking experimentally determined intracellular ionic and pH values is of considerable priority. One disadvantage with traditional EMSA analysis is that experimental environments are restricted by the conditions required for electrophoresis, whereas the relative flexibility of the SMB assay allows for more diverse conditions to be attained. The ionic strength of a solution has a substantial impact on the modes of protein-DNA interactions that can arise due to the effective neutralisation of the electrostatic attraction between protein and DNA components at higher ionic concentrations. Therefore, non-specific interactions that generally rely heavily on charge (Jen-Jacobson, 1997) are reduced but stronger specific interactions that are dependent on other factors can remain. To this extent, human p53 interactions towards sequence-specific linear dsDNA, non-sequence-specific linear dsDNA and non-specific ssDNA were assessed at low and high KCl concentrations to discriminate between specific and non-specific DNA binding. The data obtained conclusively demonstrated increasing ionic strength reduced binding to non-specific dsDNA but did not significantly reduce either sequence-specific binding or ssDNA binding. These results agreed strongly with the currently accepted models of p53-DNA binding as, although p53 does bind to dsDNA of any sequence (McKinney et al., 2004), the sequence-specific oligonucleotides investigated had previously been shown to increase p53 binding (Walter et al., 2005) and p53 is known to bind to ssDNA with affinity equal to consensus DNA (Bakalkin et al., 1994). Importantly, this data proved varied modes of DNA binding to short DNAs could be visualised without the requirement of p53 modifications, interferences with antibodies or removal of protein domains previously presumed (Anderson et al., 1997; Bessard et al., 1998; Hupp et al., 1992). The ability of p53 to remain bound to sequence-specific and ssDNA at higher KCl concentrations would almost certainly occur because of additional thermodynamic energies other than charge attraction involved in these interactions. However, studies have also shown that the presence of specific DNA molecules can act to stabilise p53 at high salt concentrations (Ishimaru et al., 2009). It is possible that this factor may also have contributed to the persevering specific-binding at higher ionic strength.

Until recently, the consensus view concerning *in vitro* p53-DNA binding analysis was that p53 required specific modification or antibody interference to stimulate precise DNA binding activity (Bessard et al., 1998). It is now accepted that unmodified p53 does exhibit efficient DNA binding properties and that modifications are not essential for discriminating between the varied modes of interaction (Kim and Deppert, 2003). The precise consequences of post-translational modifications on p53-DNA binding remain to be fully clarified now that the original notion that the CTD required modification has been disproven (McKinney et al., 2004). As modifications may influence p53 function, it was speculated that there may be differences in human p53-DNA binding between protein derived from eukaryotic (p53-I) or prokaryotic (p53-B) sources.

In vitro SMB assays were performed that were designed to differentiate between p53-B and p53-I DNA binding activity. As electrophoretic mobility analysis had identified differences in migration of these p53 proteins, it was assumed that differences in structure or sequence composition existed between both species. Binding towards consensus dsDNA, non-specific dsDNA, ssDNA and dsDNAs containing CTG repeat hairpins up to 7 repeats in size were analysed and it was found that though minor differences between both p53s could be resolved, both exhibited almost identical DNA binding activity overall. For both p53s it was identified that they bound significantly better to consensus linear DNA compared to non-specific linear dsDNA, however no significant differences were apparent when binding was compared between consensus dsDNA and ssDNA. It was also found there was no significant difference between sequence-specific binding and binding towards CTG hairpins consisting of 7 repeats. The presence of relatively small hairpin regions generally increased binding of both p53s, however highest binding was detected to larger CTG hairpins - data that was highly comparable to a previous study that used EMSA to perceive p53 binding to DNAs containing hairpins (Walter et al., 2005). As binding between both p53 species was so similar it was accepted that any differences between protein species did not significantly affect DNA binding properties.

To improve understanding of the apparent structure-dependent DNA binding activity that p53 was displaying, a myriad of different DNA molecules were produced that contained various unusual features, including mismatches, base inserts, loop regions and hairpin molecules consisting of different repeat sequences. The results collected from this analysis were consistent with currently published material as a preference towards G/A mismatches (Degtyareva et al., 2001) and higher binding to loop regions (Stros et al., 2004) compared to linear DNA of the same size were observed. It was also discovered that the size of a particular DNA molecule plays a central role in influencing non-specific p53-dsDNA binding

as from 3 non-specific dsDNAs analysed that consisted of 46 bp, 37 bp and 30 bp, p53-DNA binding was approximately 36%, 28% and 11%, respectively. Consequently, it is important during future direct comparisons of DNAs of varied structure or sequence to maintain the number of base pairs at a constant value.

Binding of p53 towards internal hairpins consisting of CTG repeats had been shown to be comparatively tighter than non-specific DNA binding and appeared similar in strength to consensus DNA binding. However, it remained unclear if this binding was exclusively structure-specific or if sequence also influenced this interaction. To address this, substrates containing internal hairpins consisting of different repeats were analysed alongside a linear dsDNA that contained 7 CTG repeats. The results from these assays established that p53 can interact strongly with hairpins independent of repeat sequence and that the structural component was integral to this activity. Importantly, binding to linear CTG repeats was comparable to p53 binding to non-specific DNA, reinforcing the probability that the hairpin structure is the overall feature determining the observed high level of binding. Slight differences in binding between the TRS hairpins may in part be explained by the alternate conformations of hairpins that result from different base pairing strengths and non-complementary base pairing. If, as the data suggests, p53 can bind to hairpins structure-dependently, it would be of interest to know what precise region of the hairpin this protein recognises and attaches to.

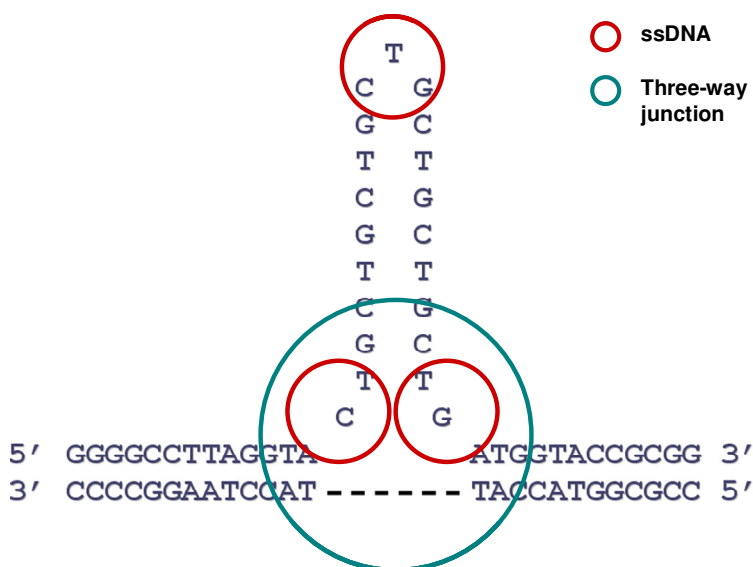


Figure 6.1 Diagram illustrating potential sites for p53 binding to [p53 (CTG)₇ HP] oligonucleotide. Sites in red indicate ssDNA regions required for formation of a central hairpin structure. ssDNA is required at the top of the hairpin to allow bending of the DNA to allow base pairing in the hairpin and fraying bases result at the base of the hairpin due to energetic stress. The region in turquoise indicates the three-way junction which is also known to be a DNA binding target site for p53.

Hairpin structures contain multiple features that have independently been found to induce p53-binding (Figure 6.1). Single-stranded regions of DNA exist both at the pinnacle of the hairpin and at the base where fraying of strands occurs as an effect of energetic stress. As shown in this study and several previous studies, ssDNA is a recognised p53 binding substrate and these regions in hairpins may be enough to be recognised and bound by p53. Alternatively, or in conjunction with ssDNA binding, the central three-way junction may also induce p53 binding. Heteroduplex joint intermediates that form during recombination are known to be recognised by p53 independent of sequence and with high affinity (Janz et al., 2002). As both ssDNA binding and three-way junction binding are mediated by the CTD, it is probable that this domain is integral to p53-hairpin binding. Another feature of certain TRS hairpins that p53 may recognise are mismatches that will induce additional conformational changes to the central duplex region consisting of repeats.

The physiological connotations of p53-hairpin interactions could be substantial as, considering the formation of secondary structures in TRS tracts present in certain human hereditary neurological disorders can be multiple-times larger than DNAs investigated in this study, the apparent affinity of p53 towards these structures would be considerable. Such activity may be an example of p53 acting as a global surveillance protein (Albrechtsen et al., 1999; Wiesmuller, 2001) and that detection of unstable TRS tracts may be an important activity of p53.

In vitro studies can determine important functions and characteristics of protein-DNA interactions; however, effective reproduction of the true intracellular environments in which these interactions occur naturally remains a distant aspiration. The introduction of synthetic molecules that mimic cellular components and achieving natural ionic concentrations and pH conditions may increase relevance of *in vitro* data, but the ability to correlate *in vitro* findings with *in vivo* findings would form a strong foundation for any hypotheses. Hence, part of this study involved an investigation into endogenous p53 DNA binding in HeLa cells using ChIP. Although levels of native p53 in HeLa cell are low as a consequence of viral E6 mediated degradation, transcriptionally active p53 is detectable (Hoppe-Seyler and Butz, 1993). Systems exist whereby levels of p53 can be over-expressed that increase detection efficiency, but unnaturally high levels of this protein has disadvantages as cellular post-translational modification machinery can become saturated and also growth arrest and apoptosis are often induced (Bessard et al., 1998).

The consequence of DNA damage upon p53-DNA interactions was of considerable interest due to the direct relationship between p53 and DNA repair (Subramanian and Griffith, 2002).

Therefore, p53-DNA binding was assessed in both UV-irradiated and non-UV-irradiated HeLa cells in order to understand if p53-mediated alterations in gene expression of p21 (regulator of the cell cycle), Mdm2 (regulator of p53) and the negative control GAPDH (glycolysis enzyme) existed. The data collected from ChIP assays did denote that a transcriptionally-active p53 was present in both UV-irradiated and non-UV irradiated HeLa cells as significantly more p21 and Mdm2 promoter DNA was precipitated than GAPDH promoter DNA. UV-irradiated HeLa cell p53 exhibited different activity as, although precipitation of p21 and Mdm2 promoter DNA remained significantly higher than GAPDH, the difference between p21 and Mdm2 was considerably reduced. Experimental variation may account for a degree of the differences observed, however p53-DNA binding may actually have been affected by the increased levels of DNA damage that would further compromise genomic integrity in the already unstable HeLa cells. As a result, an emphasis on repairing damaged DNA or the induction of apoptosis may supersede cell cycle arrest (Vousden and Lane, 2007). This would be achieved by increasing basal p53 levels and thereby increasing the probability that p53 would bind to response elements that p53 has lower affinity for. As DNA precipitation of all 3 promoter regions was increased following UV-irradiation, this may represent higher levels of intracellular p53 that would reinforce this theory.

Irrespective of this, the data does suggest p53 activity is influenced by DNA damage but it cannot be concluded that p53 sequence-specific DNA binding is increased following UV-irradiation. It appears that p53-DNA interactions in general increased following irradiation, but the underlying causes that resulted in enhanced DNA precipitation require further investigation.

6.2 Future Human p53-DNA Interaction Investigations.

The *in vitro* SMB DNA binding assay was extensively utilised during this study to determine varying properties of human p53 binding to DNA. A central part of this study was to examine the different modes of p53-DNA binding and to compare interactions with DNAs containing different structural features. As many of the results obtained were highly comparable with currently published data it would be interesting to further determine the sensitivity of the assay. In this study, the SMB assay was used to investigate sequence-specific binding to only 1 known response element, however, the recent finding of the considerable variation in affinity towards different consensus sites (Weinberg et al., 2005) would make investigating p53 binding to different consensus sequences using this assay an interesting proposition.

The SMB assay could be employed further to calculate precise affinity values for binding to clarify if these values vary between sequence-specific and structure-specific p53-DNA binding. Assays that compared binding of various p53 mutants would also be beneficial with a particular focus placed on determining the influence of the CTD on TRS interactions.

In vivo assays were particularly informative and suggested a transcriptionally active p53 was present in HeLa cells. The main aim of this work was to ultimately use a similar system to investigate p53-TRS binding, however designing a practical system that allows detection of p53 binding to authentic TRS tracts proved to be problematic. Precipitated DNA was detected using PCR whereby regions of promoter DNA approximately 100 bp in length were amplified. As TRS can consist of thousands of nucleotides of the same repeating sequence, amplification by PCR would be very difficult to perform ChIP in cells containing natural TRS. The employment of smaller TRS tracts with known flanking sequences into cells may be a practical solution to this; however it could be argued these DNAs would not reflect natural TRS tracts.

6.3 Determining Prokaryotic Ku Function.

The eukaryotic Ku70/80 heterodimer is implicated in several cellular processes that are involved in the maintenance of genomic integrity (Friedl, 2002). Perhaps the most recognised activity of Ku is as a critical component of NHEJ in the detection of DSBs and recruitment of other repair factors. Ku are also involved in the regulation of telomeres where they are postulated to prevent the potentially catastrophic joining of chromosome ends (Wang and Vasquez, 2009) and retrotransposition. The ability to bind directly to DNA is prerequisite for Ku to participate in these mechanisms and it is recognised that Ku70/Ku80 binds with high affinity towards dsDNA ends of varying structure or sequence (Falzon et al., 1993). This apparent flexibility in DNA binding is thought to arise as Ku does not directly contact DNA bases and only a small number of contacts occur with the DNA backbone (Riha et al., 2006). Ku has also been shown to bind to RNA (Yoo and Dynan, 1998), however it is commonly perceived that Ku proteins have relatively low affinity to ssDNA. Once bound, Ku does not always remain at the DNA termini but can translocate inwards, thereby allowing additional Ku proteins to attach to the same DNA molecule.

In silico investigations revealed genes homologous to the eukaryotic Ku70/Ku80 in several prokaryotic organisms that were shown to function as homodimers in two-component NHEJ systems (Della et al., 2004; Weller et al., 2002). The prokaryotic Ku homologues were found

to be considerably smaller than their eukaryotic counterparts and only generally consisted only of the central core DNA binding and dimerisation region of Ku70/Ku80. However, structural studies predicted that Ku form similar ring-like structures that shows close relation to eukaryotic Ku (Bowater and Doherty, 2006). *S. coelicolor* was found to contain 3 genes that had sequence similarity with the core region of Ku70/Ku80 and although no data existed regarding precise activity, it was hypothesised that they were involved in *S. coelicolor* NHEJ. The limited studies on DNA binding activity of characterised prokaryotic Ku proteins suggest that activity is largely identical to eukaryotic Ku, as a preference towards dsDNA ends over ssDNA or closed DNA has been documented. The apparent DNA length-dependent binding implies these proteins also possess ability to translocate inwards away from the DNA termini (Weller et al., 2002).

Preliminary analysis of recombinant versions of Ku proteins revealed unusual electrophoretic mobility in relation to the predicted molecular weights. For all 3 ScKu proteins and MtKu, migration was considerably less than expected when compared to proteins of known molecular weights in the marker. Although immunodetection against the polyhistidine tags strengthened the view that the desired proteins had been purified, analysis of the 3 ScKu proteins by MALDI-TOF protein mass fingerprinting confirmed with high confidence that these proteins were as expected. Factors that may have negatively influenced mobility include the additional histidine residues that comprise the tag or the unusually high pI values that mean during electrophoresis the higher than average positive charge of these proteins would reduce migration compared to standard proteins in the marker.

The *in vitro* SMB DNA binding assay had successfully been utilised to obtain data regarding p53-DNA binding activity that was consistent with currently accepted ideas. Therefore, this assay was again selected as a system to determine binding properties of prokaryotic NHEJ components. Assays that concentrated on comparing interactions with short dsDNA and ssDNA resolved clear differences in activity between the 3 Ku proteins. Although ScKu2 and MtKu exhibited a clear preference for dsDNA, ScKu1 and ScKu3 had vastly different activity. The apparent lack of ScKu3 binding was at first assumed to be a consequence of inactive recombinant protein or that this protein has a naturally low affinity for DNA. However, subsequent assays using plasmids demonstrated this protein does interact with DNA and that the apparent low affinity observed during the SMB assay may be a direct consequence of ScKu3 requiring a larger DNA substrate to bind. Perhaps the most interesting observation from this analysis was the consistently high ScKu1 binding to both dsDNA and ssDNA. It was found that ScKu1 bound very strongly to all oligonucleotides investigated, including hairpins and mismatched DNA (data not shown) inferring that this protein has very

high flexibility during DNA binding. Binding affinity investigations eventually revealed that ScKu1 showed a preference in binding towards dsDNA when lower concentrations of DNA were used. Therefore, it can be concluded from these experiments that ScKu1 does exhibit similar binding properties as both ScKu2 and MtKu but is able to bind more tightly to all DNAs.

Competition assays that again utilised the SMB assay produced some of the most interesting data from this study. The relative lack of effect of introducing competing DNA to Ku already bound to DNA may have been explained by Ku being difficult to remove once attached to the initial DNA molecule. However, when competing-free DNA was added simultaneously to DNA attached to SMBs it was found Ku still demonstrated much higher binding to DNA attached to SMBs. This idea was reinforced when binding of Ku to DNA-SMB remained as high as 80% even when competing DNA was 3 times in excess of DNA-SMB. Detailed analysis of assay components confirmed conditions during assays were as expected and it was concluded that there were 3 possible explanations for these observations. The first is that the Ku proteins from *S. coelicolor* are able to translocate along DNA as observed with Ku70/80 and, as the DNAs studied were only short, it may be possible that Ku could translocate and fall off the competing DNA. The second possibility is that DNA physically attached to SMBs may elicit a different conformation to free DNA that is more favourable for Ku binding. The third theory that would explain the perceived negligible effect of competing free DNA regards the potential DNA juxtaposition activity of Ku. In eukaryotes, both Ku70/80 and DNA-PKcs have been implicated in the function of synapsing 2 independent DNAs (Cary et al., 1997; DeFazio et al., 2002; Lees-Miller and Meek, 2003; Ramsden and Gellert, 1998) but this process must occur independent of DNA-PKcs in prokaryotes. To test this possibility, fluorescently labelled free DNAs were used to determine if Ku stimulated precipitation of these DNAs with SMBs. It was found that fluorescence was reduced in primary wash step samples and was higher in the final precipitated samples when Ku were present in reactions. Both of these elements suggested Ku proteins were promoting association of free DNA with DNA-SMB. It remains to be ascertained which factors were responsible for the competition assay observations, however it is possible all 3 ideas may have contributed to an extent.

It was hoped an alternative PGB *in vitro* DNA binding assay would reveal more information regarding prokaryotic Ku-DNA interactions due to the flexibility of potential DNAs that could be used as substrates. To assess similarities with SMB data, the same oligonucleotides were analysed using the PGB assay and, except for some binding by ScKu3 towards ssDNA, the results correlated highly with SMB data. When Ku-plasmid DNA binding was analysed, it was found that ScKu3 binding was very efficient to both linear and supercoiled plasmid

(activity that was closely mirrored by ScKu1), however ScKu2 binding was comparatively lower. Interestingly, ScKu1, ScKu3, MtKu and, under certain conditions ScKu2 and MtKu, all demonstrated a preference towards overhang-end DNA over blunt-end DNA. This apparent selectivity was not expected as it is generally believed Ku interactions with DNA ends is independent of sequence and structure. Potentially, regions of sequence complementarity on opposing ends of overhang-end DNA may allow these plasmids to circularise that would prevent access of DNA ends. Ku-supercoiled DNA binding was also not anticipated as Ku interactions with closed-DNAs is thought to be minimal due to the lack of DNA ends (Dynan and Yoo, 1998). However, structures such as hairpins that can be induced by supercoiling may account for increased interactions. If such structures were responsible for Ku binding, it was postulated relaxed-circular DNA would show reduced precipitation of Ku as less structures would be present. However, as almost identical results were collected it was assumed secondary structures were not wholly responsible for binding to supercoiled DNA. Therefore, from this analysis it is likely Ku can bind to DNA independent of DNA ends or hairpin structures. This would infer the currently accepted model of Ku-DNA binding of threading DNA through the central hollow channel of a Ku dimer does not account for the complete binding activity. If Ku proteins are able to spontaneously dimerise and separate in solution, theoretically, dimerisation may occur following binding of a monomer which would allow interactions with circular DNA. However, there is currently no direct evidence for this.

Alongside the detection and binding of DSBs, another characteristic of Ku is its ensuing recruitment of other repair factors required to process and ligate these dangerous lesions. For this reason it was of interest to see if prokaryotic Ku could influence the activity of T4 DNA ligase, either positively by enhancing rates of ligation (perhaps by end-alignment) or negatively by interfering with DNA ligase activity. It was found all prokaryotic Ku studied significantly altered ligation and, in the majority reactions, ligation was highly reduced or even completely prevented. Potentially, Ku that had bound and remained at the DNA termini of the plasmids may have physically blocked T4 DNA ligase access, which would explain the observed reduction in ligation. It would be assumed that, for NHEJ, such blocking would not occur as specific ligases may interact directly with Ku in a way that still allows ligation to be performed. This apparent inhibition of T4 DNA ligase activity was also documented during similar experiments with eukaryotic Ku (Blier et al., 1993; Paillard and Strauss, 1991), suggesting divergent Ku proteins induce similar effects. The recent identification that Ku can protect DNA ends (Zhu and Shuman, 2009) would provide a suitable explanation for the observed prevention of T4 DNA ligase activity as the presence of Ku at DNA termini would undoubtedly impede ligation. Interestingly, the formation of multimer plasmid species over

circularisation was also evident when certain prokaryotic Ku were present, particularly with ScKu3. This activity was identified in a previous study (Paillard and Strauss, 1991) with eukaryotic Ku and it was thought to occur following rigidification of DNA as a consequence of several internally bound Ku proteins that reduced flexibility.

Detailed *in vitro* characterisation of ScKu DNA binding properties and experiments determining the effects on T4 DNA ligase activity had been carried out, however there is currently no *in vivo* data regarding the precise functions of ScKu proteins. A section of this study has been the preparation of components for future use in the *in vivo* analysis of chromosomally-derived ScKu1 and ScKu2. Two separate approaches were taken that could eventually be used in combination to investigate multiple features of these proteins, including expression, complementation analysis, protein-protein interactions and *in vivo* DNA binding. The first approach involved the cloning of ScKu genes (including promoter and polyhistidine DNA) into specific phage-derived integration vectors, pMS82 and pSET152 that could be used to integrate these genes onto chromosomes of *S. coelicolor* strains. Expression analysis could provide useful information regarding at which lifecycle stages ScKu are expressed or if expression is increased following stresses such as UV-irradiation or exposure to high temperatures. The inclusion of polyhistidine tags on protein products would allow for immunodetection using anti-His antibodies and for direct precipitation during co-immunoprecipitation assays. Restriction digest analysis of integration vectors pMS82 and pSET152 confirmed vector identities and *sku2* was successfully amplified from cosmid DNA. However *sku1* consistently failed to amplify and attempts at cloning this gene were abandoned. Amplified *sku2* products were cloned into pCR[®]-Blunt II-TOPO vectors, sequenced to confirm inserts were as expected, and subsequently subcloned into pSET152.

The second approach was the production of polyclonal antibodies that could differentiate between ScKu1 and ScKu2 that were produced by immunising rabbits with recombinant proteins. Antibodies from final harvest bleeds displayed the most specific detection of ScKu proteins and were able to distinguish between different ScKu under appropriate conditions. These antibodies were used to investigate the presence of ScKu1 and ScKu2 in cell extracts from *E. coli* over-expressing ScKu1 and ScKu2 and bands representing relevant recombinant proteins were detected. Antibodies were also used to determine the presence of ScKu1 and ScKu2 in *S. coelicolor* M145, M145 Δ *ScKu1*, M145 Δ *ScKu2* and M145 Δ *ScKu1*/ Δ *ScKu2* cell extracts harvested at branching substrate mycelium and sporulation growth stages. The results indicated that expression of ScKu1 and ScKu2 were too low for immunodetection as no relevant bands were visible in any strain.

As the 3 ScKu proteins had been shown to interact with DNA independent of each other, it was assumed that these proteins were capable of forming homodimer complexes and so conformed to the current viewpoint regarding prokaryotic Ku (Pitcher et al., 2007). DLS data supported this concept as it was apparent that all 3 recombinant ScKu proteins were present as multiple protein species as signified by percent polydispersity values. Irrespective of this, it was postulated that it may also be possible for Ku proteins to form complexes with different Ku proteins from the same organism. Data from Co-IP assays that used anti-ScKu1 and anti-ScKu2 immunoglobulins indicated that ScKu1 and ScKu3 could interact, however ScKu2 consistently failed to precipitate any other proteins. These experiments also revealed an interaction between ScKu1 and ScLigD that may be important during the recruitment of this DNA ligase to DSBs.

For further confirmation of Co-IP findings, a separate precipitation assay was employed that used specific DNAs to precipitate ‘bait’ protein instead of immunoglobulins. This modified version of the SMB assay produced results that were in high agreement with the previous Co-IP assay as ScKu1 was found to significantly increase precipitation of ScKu3 and ScLigD when appropriate DNAs were present. Unlike antibody-dependent Co-IP results, there was a suggestion that ScKu2 was also capable of interacting with ScKu1 and ScKu3, however these interactions were generally weaker and inconsistent so may have been experimental anomalies. The discovery that an interaction exists between ScKu1 and ScKu3 was supported by *in silico* investigations as 2 out of 3 alternate *Streptomyces* species contained these 2 proteins, but none contained ScKu2. This data implies recently *S. coelicolor* has acquired the *scku2* gene, possibly by horizontal gene transfer following its evolutionary divergence away from the other *Streptomyces* species. This would explain the lack of interaction with ScKu1 and ScKu3 and the observation that this protein also did not interact with ScLigD suggests this protein may not be implicated in *S. coelicolor* NHEJ.

6.4 Future ScKu NHEJ Investigations.

A significant area of this study involved the analysis of DNA binding characteristics of prokaryotic Ku proteins using optimised versions of recently developed *in vitro* assays. Several properties regarding interactions with different DNAs have been identified, however these assays could be employed further to determine more accurate values regarding binding affinity. Although this was attempted during this study, the data collected was not sufficient to calculate precise values that would be attained by investigating an increased number of lower DNA concentrations. Another important aspect of calculating Ku binding affinity

would be to ascertain the number of Ku that bind to a single DNA molecule of particular size – a feature that should be possible to determine using AUC experiments.

One other area of this study that would be beneficial to explore further is the apparent DNA substrate synapsis activity that ScKu1 and ScKu2 proteins exhibited. Assays investigating the precipitation of fluorescently labelled free DNAs by SMB-DNA-Ku complexes suggested a degree of alignment of DNAs was occurring, however the data was not as conclusive as previous competition assays suggested. Repeats of similar assays but with the use of fluorescently-labelled dsDNA rather than ssDNA would be a suitable starting point to determine if synapsis efficiency increases with DNAs shown to induce higher levels of Ku binding. Alternatively, atomic force microscopy (AFM) could be used for direct visualisation of any DNA bridging activity that prokaryotic Ku possess. AFM can produce high resolution images of protein-nucleic acid complexes (Cary et al., 1997) and, therefore, would be beneficial in exploring this potential Ku function.

A separate function of prokaryotic Ku that has recently been documented (Zhu and Shuman, 2009) is the ability to protect DSB ends from exonuclease degradation. A relatively simple assay could be developed where Ku proteins are initially incubated with linear DNAs followed by the addition of exonucleases. Comparisons of degradation of DNAs without Ku addition would establish if Ku can protect DNA from enzymatic digestion and this would also indicate if Ku remain at the ends of DNA (as indicated during *in vitro* T4 DNA Ligase assays) as well as translocating inwards.

Perhaps the most valuable *in vitro* achievement regarding *S. coelicolor* NHEJ would be the accurate replication of the entire *S. coelicolor* NHEJ apparatus capable of aligning and ligating 2 DNA molecules. So far, several components have been identified aside from ScKu proteins including 3 primases, 2 NAD⁺-dependent DNA ligases, 3 ATP-dependent DNA ligases and 1 nuclease (Sayer, 2008). As recombinant versions of these proteins are already available, the next steps would be to identify appropriate conditions in which protein activity is optimal to allow efficient DNA end joining activity.

The cloning of recombinant *scku2* into integration vector pSET152 and the production of antibodies specific for ScKu1 and ScKu2 has provided a platform for several potential *in vivo* investigations to examine different features of ScKu proteins. The presence of the native *scku2* promoter will allow natural expression levels to be determined during the various life-stages of *S. coelicolor* and to establish if levels increase following induction of DNA damage or under various stressful conditions. Alternatively, gene expression could be monitored

using reportable fusion markers inserted adjacent to various *S. coelicolor* genes associated with NHEJ that, alongside expression trends, would also aid identification of the localisation of proteins via fluorescent marker tracking.

As previously mentioned, specific ScKu antibodies provide the opportunity to determine *in vivo* protein-protein interactions via Co-IP assays. Due to the low cross-reactivity exhibited by FHB antibodies, appropriate antibodies may be used to precipitate a specific ScKu protein followed by immunodetection with alternate ScKu antibodies to examine if interactions between different ScKu proteins occur. This analysis could provide further information regarding dimerisation of ScKu proteins.

The pivotal role Ku plays during maintenance of eukaryotic telomeres is well established (Riha et al., 2006) where Ku proteins are believed to be instrumental in preventing the potentially catastrophic joining of chromosome ends. No evidence exists that implicates prokaryotic Ku in similar maintenance of genomes but, as most prokaryotes do not have linear chromosomes, the protection of DNA termini is not a problem for these organisms. However, *S. coelicolor* does contain a linear genome and also linear plasmids that are capable of circularisation (Bao and Cohen, 2001). As with eukaryotic chromosome ends, the chromosome of *S. coelicolor* also contains telomeres consisting of inverted repeat sequences with the 5' end covalently attached to terminal proteins. Both Tap and Tpg proteins have been identified as important components in the maintenance of *S. coelicolor* telomeres to date (Bao and Cohen, 2003), however it would not be surprising if ScKu were also involved in regulating these structures. To this extent, specific ScKu antibodies could be used to perform *in vivo* ChIP assays to identify if ScKu directly interact with telomeric regions.

The maintenance of genomic integrity is fundamental to cellular life and understanding the precise interactions that govern this stability is central to preventing the events that can lead to disease. The complexity of macromolecular interactions requires a varied arsenal of techniques that can be used in conjunction to fully characterise how these mechanisms work. This study has used several different biochemical assays to resolve the activities of DNA repair proteins believed to recognise dangers to genomic stability. The flexibility of these assays has provided opportunities for methodical adaptations so that additional functions have been resolved using similar systems. The frequent high correlation with published data and high reproducibility indicate these assays would be appropriate for future protein investigations.

References

- Ahn, J. and Prives, C. (2001) The C-terminus of p53: the more you learn the less you know. *Nat Struct Biol*, **8**, 730-732.
- Albrechtsen, N., Dornreiter, I., Grosse, F., Kim, E., Wiesmuller, L. and Deppert, W. (1999) Maintenance of genomic integrity by p53: complementary roles for activated and non-activated p53. *Oncogene*, **18**, 7706-7717.
- Altona, C., Pikkemaat, J.A. and Overmars, F.J. (1996) Three-way and four-way junctions in DNA: a conformational viewpoint. *Curr Opin Struct Biol*, **6**, 305-316.
- Anderson, M.E., Woelker, B., Reed, M., Wang, P. and Tegtmeyer, P. (1997) Reciprocal interference between the sequence-specific core and nonspecific C-terminal DNA binding domains of p53: implications for regulation. *Mol Cell Biol*, **17**, 6255-6264.
- Angert, E.R. (2005) Alternatives to binary fission in bacteria. *Nat Rev Microbiol*, **3**, 214-224.
- Appella, E. and Anderson, C.W. (2001) Post-translational modifications and activation of p53 by genotoxic stresses. *Eur J Biochem*, **268**, 2764-2772.
- Aravind, L. and Koonin, E.V. (2000) SAP - a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem Sci*, **25**, 112-114.
- Aravind, L. and Koonin, E.V. (2001) Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res*, **11**, 1365-1374.
- Arrowsmith, C.H. (1999) Structure and function in the p53 family. *Cell Death Differ*, **6**, 1169-1173.
- Bacolla, A., Jaworski, A., Larson, J.E., Jakupciak, J.P., Chuzhanova, N., Abeyasinghe, S.S., O'Connell, C.D., Cooper, D.N. and Wells, R.D. (2004) Breakpoints of gross deletions coincide with non-B DNA conformations. *Proc Natl Acad Sci U S A*, **101**, 14162-14167.
- Bacolla, A. and Wells, R.D. (2004) Non-B DNA conformations, genomic rearrangements, and human disease. *J Biol Chem*, **279**, 47411-47414.
- Bacolla, A. and Wells, R.D. (2009) Non-B DNA conformations as determinants of mutagenesis and human disease. *Mol Carcinog*, **48**, 273-285.
- Bakalkin, G., Selivanova, G., Yakovleva, T., Kiseleva, E., Kashuba, E., Magnusson, K.P., Szekely, L., Klein, G., Terenius, L. and Wiman, K.G. (1995) p53 binds single-stranded DNA ends through the C-terminal domain and internal DNA segments via the middle domain. *Nucleic Acids Res*, **23**, 362-369.
- Bakalkin, G., Yakovleva, T., Selivanova, G., Magnusson, K.P., Szekely, L., Kiseleva, E., Klein, G., Terenius, L. and Wiman, K.G. (1994) p53 binds single-stranded DNA ends and catalyzes DNA renaturation and strand transfer. *Proc Natl Acad Sci U S A*, **91**, 413-417.
- Bao, K. and Cohen, S.N. (2001) Terminal proteins essential for the replication of linear plasmids and chromosomes in *Streptomyces*. *Genes Dev*, **15**, 1518-1527.
- Bao, K. and Cohen, S.N. (2003) Recruitment of terminal protein to the ends of *Streptomyces* linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication. *Genes Dev*, **17**, 774-785.
- Bargonetti, J. and Manfredi, J.J. (2002) Multiple roles of the tumor suppressor p53. *Curr Opin Oncol*, **14**, 86-91.
- Barlow, J.H., Lisby, M. and Rothstein, R. (2008) Differential regulation of the cellular response to DNA double-strand breaks in G1. *Mol Cell*, **30**, 73-85.
- Bates, A.D. and Maxwell, A. (2005) DNA Topology. *Oxford University Press*.
- Batra, J., Xu, K. and Zhou, H.X. (2009) Nonadditive effects of mixed crowding on protein stability. *Proteins*.

- Bayle, J.H., Elenbaas, B. and Levine, A.J. (1995) The carboxyl-terminal domain of the p53 protein regulates sequence-specific DNA binding through its nonspecific nucleic acid-binding activity. *Proc Natl Acad Sci U S A*, **92**, 5729-5733.
- Becker, W.M., Kleinsmith, L.J., Hardin, J. and Berton, G.P. (2009) *The World of the Cell*. Benjamin Cummings, Seventh Edition.
- Bellamy, S.R., Milsom, S.E., Scott, D.J., Daniels, L.E., Wilson, G.G. and Halford, S.E. (2005) Cleavage of individual DNA strands by the different subunits of the heterodimeric restriction endonuclease BbvCI. *J Mol Biol*, **348**, 641-653.
- Bernstein, K.A. and Rothstein, R. (2009) At loose ends: resecting a double-strand break. *Cell*, **137**, 807-810.
- Berti, P.J. and McCann, J.A. (2006) Toward a detailed understanding of base excision repair enzymes: transition state and mechanistic analyses of N-glycoside hydrolysis and N-glycoside transfer. *Chem Rev*, **106**, 506-555.
- Bessard, A.C., Garay, E., Lacronique, V., Legros, Y., Demarquay, C., Houque, A., Portefaix, J.M., Granier, C. and Soussi, T. (1998) Regulation of the specific DNA binding activity of *Xenopus laevis* p53: evidence for conserved regulation through the carboxy-terminus of the protein. *Oncogene*, **16**, 883-890.
- Blier, P.R., Griffith, A.J., Craft, J. and Hardin, J.A. (1993) Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *J Biol Chem*, **268**, 7594-7601.
- Boehme, K.A. and Blattner, C. (2009) Regulation of p53--insights into a complex process. *Crit Rev Biochem Mol Biol*, **44**, 367-392.
- Bowater, R. and Doherty, A.J. (2006) Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet*, **2**, e8.
- Bowater, R.P. (2005) Supercoiled DNA: Structure. *Encyclopedia of Life Sciences*.
- Bowater, R.P. and Wells, R.D. (2001) The intrinsically unstable life of DNA triplet repeats associated with human hereditary disorders. *Prog Nucleic Acid Res Mol Biol*, **66**, 159-202.
- Brachmann, R.K. (2004) p53 mutants: the achilles' heel of human cancers? *Cell Cycle*, **3**, 1030-1034.
- Bullock, A.N., Henckel, J. and Fersht, A.R. (2000) Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene*, **19**, 1245-1256.
- Cain, C., Miller, S., Ahn, J. and Prives, C. (2000) The N terminus of p53 regulates its dissociation from DNA. *J Biol Chem*, **275**, 39944-39953.
- Calladine, C.R. (1982) Mechanics of sequence-dependent stacking of bases in B-DNA. *J Mol Biol*, **161**, 343-352.
- Cary, R.B., Peterson, S.R., Wang, J., Bear, D.G., Bradbury, E.M. and Chen, D.J. (1997) DNA looping by Ku and the DNA-dependent protein kinase. *Proc Natl Acad Sci U S A*, **94**, 4267-4272.
- Chakrabarti, R. and Schutt, C.E. (2001) The enhancement of PCR amplification by low molecular weight amides. *Nucleic Acids Res*, **29**, 2377-2381.
- Chastain, P.D., 2nd, Eichler, E.E., Kang, S., Nelson, D.L., Levene, S.D. and Sinden, R.R. (1995) Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. *Biochemistry*, **34**, 16125-16131.
- Christmann, M., Tomicic, M.T., Roos, W.P. and Kaina, B. (2003) Mechanisms of human DNA repair: an update. *Toxicology*, **193**, 3-34.
- Collins, J., Volckaert, G. and Nevers, P. (1982) Precise and nearly-precise excision of the symmetrical inverted repeats of Tn5; common features of recA-independent deletion events in *Escherichia coli*. *Gene*, **19**, 139-146.
- Combes, P., Till, R., Bee, S. and Smith, M.C. (2002) The streptomyces genome contains multiple pseudo-attB sites for the (phi)C31-encoded site-specific recombination system. *J Bacteriol*, **184**, 5746-5752.
- Cooper, J.P. and Hagerman, P.J. (1987) Gel electrophoretic analysis of the geometry of a DNA four-way junction. *J Mol Biol*, **198**, 711-719.

- Coulocheri, S.A., Pigis, D.G., Papavassiliou, K.A. and Papavassiliou, A.G. (2007) Hydrogen bonds in protein-DNA complexes: where geometry meets plasticity. *Biochimie*, **89**, 1291-1303.
- Critchlow, S.E. and Jackson, S.P. (1998) DNA end-joining: from yeast to man. *Trends Biochem Sci*, **23**, 394-398.
- Croy, J.E. and Wuttke, D.S. (2006) Themes in ssDNA recognition by telomere-end protection proteins. *Trends Biochem Sci*, **31**, 516-525.
- Curtin, J.C. and Spinella, M.J. (2005) p53 in human embryonal carcinoma: identification of a transferable, transcriptional repression domain in the N-terminal region of p53. *Oncogene*, **24**, 1481-1490.
- d'Adda di Fagagna, F., Weller, G.R., Doherty, A.J. and Jackson, S.P. (2003) The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep*, **4**, 47-52.
- Daley, J.M., Palmbo, P.L., Wu, D. and Wilson, T.E. (2005) Nonhomologous end joining in yeast. *Annu Rev Genet*, **39**, 431-451.
- Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, **97**, 6640-6645.
- Dawson, R., Muller, L., Dehner, A., Klein, C., Kessler, H. and Buchner, J. (2003) The N-terminal domain of p53 is natively unfolded. *J Mol Biol*, **332**, 1131-1141.
- de Vries, E., van Driel, W., Bergsma, W.G., Arnberg, A.C. and van der Vliet, P.C. (1989) HeLa nuclear protein recognizing DNA termini and translocating on DNA forming a regular DNA-multimeric protein complex. *J Mol Biol*, **208**, 65-78.
- DeFazio, L.G., Stansel, R.M., Griffith, J.D. and Chu, G. (2002) Synapsis of DNA ends by DNA-dependent protein kinase. *Embo J*, **21**, 3192-3200.
- Degtyareva, N., Subramanian, D. and Griffith, J.D. (2001) Analysis of the binding of p53 to DNAs containing mismatched and bulged bases. *J Biol Chem*, **276**, 8778-8784.
- Dehner, A., Klein, C., Hansen, S., Muller, L., Buchner, J., Schwaiger, M. and Kessler, H. (2005) Cooperative binding of p53 to DNA: regulation by protein-protein interactions through a double salt bridge. *Angew Chem Int Ed Engl*, **44**, 5247-5251.
- Della, M., Palmbo, P.L., Tseng, H.M., Tonkin, L.M., Daley, J.M., Topper, L.M., Pitcher, R.S., Tomkinson, A.E., Wilson, T.E. and Doherty, A.J. (2004) Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science*, **306**, 683-685.
- Doherty, A.J. and Jackson, S.P. (2001) DNA repair: how Ku makes ends meet. *Curr Biol*, **11**, R920-924.
- Doherty, A.J., Jackson, S.P. and Weller, G.R. (2001) Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Lett*, **500**, 186-188.
- Dominy, B.N., Perl, D., Schmid, F.X. and Brooks, C.L., 3rd. (2002) The effects of ionic strength on protein stability: the cold shock protein family. *J Mol Biol*, **319**, 541-554.
- Duckett, D.R., Murchie, A.I., Diekmann, S., von Kitzing, E., Kemper, B. and Lilley, D.M. (1988) The structure of the Holliday junction, and its resolution. *Cell*, **55**, 79-89.
- Dudenhoffer, C., Rohaly, G., Will, K., Deppert, W. and Wiesmuller, L. (1998) Specific mismatch recognition in heteroduplex intermediates by p53 suggests a role in fidelity control of homologous recombination. *Mol Cell Biol*, **18**, 5332-5342.
- Dynan, W.S. and Yoo, S. (1998) Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Res*, **26**, 1551-1559.
- el-Deiry, W.S., Tokino, T., Velculescu, V.E., Levy, D.B., Parsons, R., Trent, J.M., Lin, D., Mercer, W.E., Kinzler, K.W. and Vogelstein, B. (1993) WAF1, a potential mediator of p53 tumor suppression. *Cell*, **75**, 817-825.
- Ensign, J.C. (1978) Formation, properties, and germination of actinomycete spores. *Annu Rev Microbiol*, **32**, 185-219.
- Espinosa, J.M. and Emerson, B.M. (2001) Transcriptional regulation by p53 through intrinsic DNA/chromatin binding and site-directed cofactor recruitment. *Mol Cell*, **8**, 57-69.
- Falzon, M., Fewell, J.W. and Kuff, E.L. (1993) EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single- to double-strand transitions in DNA. *J Biol Chem*, **268**, 10546-10552.

- Fleck, O. and Nielsen, O. (2004) DNA repair. *J Cell Sci*, **117**, 515-517.
- Fridman, J.S. and Lowe, S.W. (2003) Control of apoptosis by p53. *Oncogene*, **22**, 9030-9040.
- Friedl, A.A. (2002) Ku and the Stability of the Genome. *J Biomed Biotechnol*, **2**, 61-65.
- Gatchel, J.R. and Zoghbi, H.Y. (2005) Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet*, **6**, 743-755.
- Gehring, A.M., Wang, S.T., Kearns, D.B., Storer, N.Y. and Losick, R. (2004) Novel genes that influence development in *Streptomyces coelicolor*. *J Bacteriol*, **186**, 3570-3577.
- Genschel, J., Bazemore, L.R. and Modrich, P. (2002) Human exonuclease I is required for 5' and 3' mismatch repair. *J Biol Chem*, **277**, 13302-13311.
- Godar, S., Ince, T.A., Bell, G.W., Feldser, D., Donaher, J.L., Bergh, J., Liu, A., Miu, K., Watnick, R.S., Reinhardt, F., McAllister, S.S., Jacks, T. and Weinberg, R.A. (2008) Growth-inhibitory and tumor-suppressive functions of p53 depend on its repression of CD44 expression. *Cell*, **134**, 62-73.
- Gohler, T., Jager, S., Warnecke, G., Yasuda, H., Kim, E. and Deppert, W. (2005) Mutant p53 proteins bind DNA in a DNA structure-selective mode. *Nucleic Acids Res*, **33**, 1087-1100.
- Gohler, T., Reimann, M., Cherny, D., Walter, K., Warnecke, G., Kim, E. and Deppert, W. (2002) Specific interaction of p53 with target binding sites is determined by DNA conformation and is regulated by the C-terminal domain. *J Biol Chem*, **277**, 41192-41203.
- Gowers, D.M., Wilson, G.G. and Halford, S.E. (2005) Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proc Natl Acad Sci U S A*, **102**, 15883-15888.
- Gregory, M.A., Till, R. and Smith, M.C. (2003) Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors. *J Bacteriol*, **185**, 5320-5323.
- Hakem, R. (2008) DNA-damage repair; the good, the bad, and the ugly. *Embo J*, **27**, 589-605.
- Halford, S.E. (2009) An end to 40 years of mistakes in DNA-protein association kinetics? *Biochem Soc Trans*, **37**, 343-348.
- Harms, K., Nozell, S. and Chen, X. (2004) The common and distinct target genes of the p53 family transcription factors. *Cell Mol Life Sci*, **61**, 822-842.
- Hartl, F.U. and Hayer-Hartl, M. (2009) Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol*, **16**, 574-581.
- Hashida, H., Goto, J., Kurisaki, H., Mizusawa, H. and Kanazawa, I. (1997) Brain regional differences in the expansion of a CAG repeat in the spinocerebellar ataxias: dentatorubral-pallidoluysian atrophy, Machado-Joseph disease, and spinocerebellar ataxia type 1. *Ann Neurol*, **41**, 505-511.
- Heidenreich, E., Novotny, R., Kneidinger, B., Holzmann, V. and Wintersberger, U. (2003) Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *Embo J*, **22**, 2274-2283.
- Heiter, D.F., Lunnen, K.D. and Wilson, G.G. (2005) Site-specific DNA-nicking mutants of the heterodimeric restriction endonuclease R.BbvCI. *J Mol Biol*, **348**, 631-640.
- Hirochika, H. and Sakaguchi, K. (1982) Analysis of linear plasmids isolated from *Streptomyces*: association of protein with the ends of the plasmid DNA. *Plasmid*, **7**, 59-65.
- Hirsch, C.F. and Ensign, J.C. (1976) Nutritionally defined conditions for germination of *Streptomyces viridochromogenes* spores. *J Bacteriol*, **126**, 13-23.
- Ho, W.C., Fitzgerald, M.X. and Marmorstein, R. (2006) Structure of the p53 core domain dimer bound to DNA. *J Biol Chem*, **281**, 20494-20502.
- Hoeijmakers, J.H. (2001) Genome maintenance mechanisms for preventing cancer. *Nature*, **411**, 366-374.
- Holmberg, A., Blomstergren, A., Nord, O., Lukacs, M., Lundeberg, J. and Uhlen, M. (2005) The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. *Electrophoresis*, **26**, 501-510.

- Hoppe-Seyler, F. and Butz, K. (1993) Repression of endogenous p53 transactivation function in HeLa cervical carcinoma cells by human papillomavirus type 16 E6, human mdm-2, and mutant p53. *J Virol*, **67**, 3111-3117.
- Hopwood, D.A., Lydiate, D.J., Malpartida, F. and Wright, H.M. (1985) Conjugative sex plasmids of *Streptomyces*. *Basic Life Sci*, **30**, 615-634.
- Horn, H.F. and Vousden, K.H. (2007) Coping with stress: multiple ways to activate p53. *Oncogene*, **26**, 1306-1316.
- Hupp, T.R. and Lane, D.P. (1994) Allosteric activation of latent p53 tetramers. *Curr Biol*, **4**, 865-875.
- Hupp, T.R., Meek, D.W., Midgley, C.A. and Lane, D.P. (1992) Regulation of the specific DNA binding function of p53. *Cell*, **71**, 875-886.
- Hutchings, M.I., Hong, H.J., Leibovitz, E., Sutcliffe, I.C. and Buttner, M.J. (2006) The sigma(E) cell envelope stress response of *Streptomyces coelicolor* is influenced by a novel lipoprotein, CseA. *J Bacteriol*, **188**, 7222-7229.
- Inga, A., Storici, F., Darden, T.A. and Resnick, M.A. (2002) Differential transactivation by the p53 transcription factor is highly dependent on p53 level and promoter target sequence. *Mol Cell Biol*, **22**, 8612-8625.
- Ip, S.C., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M. and West, S.C. (2008) Identification of Holliday junction resolvases from humans and yeast. *Nature*, **456**, 357-361.
- Ishimaru, D., Ano Bom, A.P., Lima, L.M., Quesado, P.A., Oyama, M.F., de Moura Gallo, C.V., Cordeiro, Y. and Silva, J.L. (2009) Cognate DNA stabilizes the tumor suppressor p53 and prevents misfolding and aggregation. *Biochemistry*, **48**, 6126-6135.
- Itahana, K., Dimri, G. and Campisi, J. (2001) Regulation of cellular senescence by p53. *Eur J Biochem*, **268**, 2784-2791.
- Iyer, R.R., Pluciennik, A., Burdett, V. and Modrich, P.L. (2006) DNA mismatch repair: functions and mechanisms. *Chem Rev*, **106**, 302-323.
- Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071-1078.
- Janz, C., Susse, S. and Wiesmuller, L. (2002) p53 and recombination intermediates: role of tetramerization at DNA junctions in complex formation and exonucleolytic degradation. *Oncogene*, **21**, 2130-2140.
- Jen-Jacobson, L. (1997) Protein-DNA recognition complexes: conservation of structure and binding energy in the transition state. *Biopolymers*, **44**, 153-180.
- Joerger, A.C. and Fersht, A.R. (2008) Structural biology of the tumor suppressor p53. *Annu Rev Biochem*, **77**, 557-582.
- Juarez, R., Ruiz, J.F., Nick McElhinny, S.A., Ramsden, D. and Blanco, L. (2006) A specific loop in human DNA polymerase mu allows switching between creative and DNA-instructed synthesis. *Nucleic Acids Res*, **34**, 4572-4582.
- Jun, S.H., Kim, T.G. and Ban, C. (2006) DNA mismatch repair system. Classical and fresh roles. *Febs J*, **273**, 1609-1619.
- Kabata, H., Kurosawa, O., Arai, I., Washizu, M., Margaron, S.A., Glass, R.E. and Shimamoto, N. (1993) Visualization of single molecules of RNA polymerase sliding along DNA. *Science*, **262**, 1561-1563.
- Kaczmariski, W. and Khan, S.A. (1993) Lupus autoantigen Ku protein binds HIV-1 TAR RNA in vitro. *Biochem Biophys Res Commun*, **196**, 935-942.
- Kaesler, M.D. and Iggo, R.D. (2002) Chromatin immunoprecipitation analysis fails to support the latency model for regulation of p53 DNA binding activity in vivo. *Proc Natl Acad Sci U S A*, **99**, 95-100.
- Kahn, J.D., Yun, E. and Crothers, D.M. (1994) Detection of localized DNA flexibility. *Nature*, **368**, 163-166.
- Kampmann, M. (2004) Obstacle bypass in protein motion along DNA by two-dimensional rather than one-dimensional sliding. *J Biol Chem*, **279**, 38715-38720.

- Kao-Huang, Y., Revzin, A., Butler, A.P., O'Conner, P., Noble, D.W. and von Hippel, P.H. (1977) Nonspecific DNA binding of genome-regulating proteins as a biological control mechanism: measurement of DNA-bound *Escherichia coli* lac repressor in vivo. *Proc Natl Acad Sci U S A*, **74**, 4228-4232.
- Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000) Practical *Streptomyces* genetics. Norwich, Crowes.
- Kim, E., Albrechtsen, N. and Deppert, W. (1997) DNA-conformation is an important determinant of sequence-specific DNA binding by tumor suppressor p53. *Oncogene*, **15**, 857-869.
- Kim, E. and Deppert, W. (2003) The complex interactions of p53 with target DNA: we learn as we go. *Biochem Cell Biol*, **81**, 141-150.
- Kim, E. and Deppert, W. (2006) The versatile interactions of p53 with DNA: when flexibility serves specificity. *Cell Death Differ*, **13**, 885-889.
- Kim, S.T., Lim, D.S., Canman, C.E. and Kastan, M.B. (1999) Substrate specificities and identification of putative substrates of ATM kinase family members. *J Biol Chem*, **274**, 37538-37543.
- Kinashi, H. and Shimaji, M. (1987) Detection of giant linear plasmids in antibiotic producing strains of *Streptomyces* by the OFAGE technique. *J Antibiot (Tokyo)*, **40**, 913-916.
- Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E. and Shakked, Z. (2006) Structural basis of DNA recognition by p53 tetramers. *Mol Cell*, **22**, 741-753.
- Kobayashi, H., Simmons, L.A., Yuan, D.S., Broughton, W.J. and Walker, G.C. (2008) Multiple Ku orthologues mediate DNA non-homologous end-joining in the free-living form and during chronic infection of *Sinorhizobium meliloti*. *Mol Microbiol*, **67**, 350-363.
- Kohn, W.D., Kay, C.M. and Hodges, R.S. (1997) Salt effects on protein stability: two-stranded alpha-helical coiled-coils containing inter- or intrahelical ion pairs. *J Mol Biol*, **267**, 1039-1052.
- Kovtun, I.V. and McMurray, C.T. (2008) Features of trinucleotide repeat instability in vivo. *Cell Res*, **18**, 198-213.
- Kroll, J. (2002) Molecular chaperones and the process of cellular immortalization in vitro. *Biogerontology*, **3**, 183-185.
- Kulkarni, A. and Wilson, D.M., 3rd. (2008) The involvement of DNA-damage and -repair defects in neurological dysfunction. *Am J Hum Genet*, **82**, 539-566.
- Kur, J., Olszewski, M., Dlugolecka, A. and Filipkowski, P. (2005) Single-stranded DNA-binding proteins (SSBs) -- sources and applications in molecular biology. *Acta Biochim Pol*, **52**, 569-574.
- Kysela, B., Doherty, A.J., Chovanec, M., Stiff, T., Ameer-Beg, S.M., Vojnovic, B., Girard, P.M. and Jeggo, P.A. (2003) Ku stimulation of DNA ligase IV-dependent ligation requires inward movement along the DNA molecule. *J Biol Chem*, **278**, 22466-22474.
- Laptenko, O. and Prives, C. (2006) Transcriptional regulation by p53: one protein, many possibilities. *Cell Death Differ*, **13**, 951-961.
- Lee, S., Elenbaas, B., Levine, A. and Griffith, J. (1995) p53 and its 14 kDa C-terminal domain recognize primary DNA damage in the form of insertion/deletion mismatches. *Cell*, **81**, 1013-1020.
- Lee, S.T. and Kim, M. (2006) Aging and neurodegeneration. Molecular mechanisms of neuronal loss in Huntington's disease. *Mech Ageing Dev*, **127**, 432-435.
- Leeflang, E.P., Tavaré, S., Marjoram, P., Neal, C.O., Srinidhi, J., MacFarlane, H., MacDonald, M.E., Gusella, J.F., de Young, M., Wexler, N.S. and Arnheim, N. (1999) Analysis of germline mutation spectra at the Huntington's disease locus supports a mitotic mutation mechanism. *Hum Mol Genet*, **8**, 173-183.
- Lees-Miller, S.P. and Meek, K. (2003) Repair of DNA double strand breaks by non-homologous end joining. *Biochimie*, **85**, 1161-1173.

- Lehman, J.A., Hoelz, D.J. and Turchi, J.J. (2008) DNA-Dependent Conformational Changes in the Ku Heterodimer. *Biochemistry*.
- Li, C., Lin, M. and Liu, J. (2004) Identification of PRC1 as the p53 target gene uncovers a novel function of p53 in the regulation of cytokinesis. *Oncogene*, **23**, 9336-9347.
- Li, G., Nelsen, C. and Hendrickson, E.A. (2002) Ku86 is essential in human somatic cells. *Proc Natl Acad Sci U S A*, **99**, 832-837.
- Liang, F. and Jasin, M. (1996) Ku80-deficient cells exhibit excess degradation of extrachromosomal DNA. *J Biol Chem*, **271**, 14405-14411.
- Lilley, D.M. (1997) All change at Holliday junction. *Proc Natl Acad Sci U S A*, **94**, 9513-9515.
- Liu, Y. and Kulesz-Martin, M. (2001) p53 protein at the hub of cellular DNA damage response pathways through sequence-specific and non-sequence-specific DNA binding. *Carcinogenesis*, **22**, 851-860.
- Lombard, D.B., Chua, K.F., Mostoslavsky, R., Franco, S., Gostissa, M. and Alt, F.W. (2005) DNA repair, genome stability, and aging. *Cell*, **120**, 497-512.
- Longley, M.J., Pierce, A.J. and Modrich, P. (1997) DNA polymerase delta is required for human mismatch repair in vitro. *J Biol Chem*, **272**, 10917-10921.
- Macville, M., Schrock, E., Padilla-Nash, H., Keck, C., Ghadimi, B.M., Zimonjic, D., Popescu, N. and Ried, T. (1999) Comprehensive and definitive molecular cytogenetic characterization of HeLa cells by spectral karyotyping. *Cancer Res*, **59**, 141-150.
- Malyarchuk, S., Wright, D., Castore, R., Klepper, E., Weiss, B., Doherty, A.J. and Harrison, L. (2007) Expression of Mycobacterium tuberculosis Ku and Ligase D in Escherichia coli results in RecA and RecB-independent DNA end-joining at regions of microhomology. *DNA Repair (Amst)*, **6**, 1413-1424.
- Marcel, V. and Hainaut, P. (2009) p53 isoforms - a conspiracy to kidnap p53 tumor suppressor activity? *Cell Mol Life Sci*, **66**, 391-406.
- McConnell, B. and von Hippel, P.H. (1970) Hydrogen exchange as a probe of the dynamic structure of DNA. I. General acid-base catalysis. *J Mol Biol*, **50**, 297-316.
- McKinney, K., Mattia, M., Gottifredi, V. and Prives, C. (2004) p53 linear diffusion along DNA requires its C terminus. *Mol Cell*, **16**, 413-424.
- McKinney, K. and Prives, C. (2002) Efficient specific DNA binding by p53 requires both its central and C-terminal domains as revealed by studies with high-mobility group 1 protein. *Mol Cell Biol*, **22**, 6797-6808.
- Meletis, K., Wirta, V., Hede, S.M., Nister, M., Lundeberg, J. and Frisen, J. (2006) p53 suppresses the self-renewal of adult neural stem cells. *Development*, **133**, 363-369.
- Migueluez, E.M., Hardisson, C. and Manzanal, M.B. (2000) Streptomyces: a new model to study cell death. *Int Microbiol*, **3**, 153-158.
- Mimori, T., Akizuki, M., Yamagata, H., Inada, S., Yoshida, S. and Homma, M. (1981) Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *J Clin Invest*, **68**, 611-620.
- Mirkin, S.M. (2006) DNA structures, repeat expansions and human hereditary disorders. *Curr Opin Struct Biol*, **16**, 351-358.
- Mirkin, S.M. (2007) Expandable DNA repeats and human disease. *Nature*, **447**, 932-940.
- Moffatt, B.A. and Studier, F.W. (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell*, **49**, 221-227.
- Murray-Zmijewski, F., Slee, E.A. and Lu, X. (2008) A complex barcode underlies the heterogeneous response of p53 to stress. *Nat Rev Mol Cell Biol*, **9**, 702-712.
- Myers, R.H. (2004) Huntington's disease genetics. *NeuroRx*, **1**, 255-262.
- N'Soukpoe-Kossi, C.N., Diamantoglou, S. and Tajmir-Riahi, H.A. (2008) DNase I - DNA interaction alters DNA and protein conformations. *Biochem Cell Biol*, **86**, 244-250.
- Nelson, D.L. and Cox, M.M. (2000) Lehninger Principles of Biochemistry. Worth Publishers, Third Edition.
- Oda, M. and Nakamura, H. (2000) Thermodynamic and kinetic analyses for understanding sequence-specific DNA recognition. *Genes Cells*, **5**, 319-326.

- Oh, D.B., Kim, Y.G. and Rich, A. (2002) Z-DNA-binding proteins can act as potent effectors of gene expression in vivo. *Proc Natl Acad Sci U S A*, **99**, 16666-16671.
- Okorokov, A.L. and Orlova, E.V. (2009) Structural biology of the p53 tumour suppressor. *Curr Opin Struct Biol*, **19**, 197-202.
- Oren, M. (2003) Decision making by p53: life, death and cancer. *Cell Death Differ*, **10**, 431-442.
- Pabo, C.O. and Nekludova, L. (2000) Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition? *J Mol Biol*, **301**, 597-624.
- Paillard, S. and Strauss, F. (1991) Analysis of the mechanism of interaction of simian Ku protein with DNA. *Nucleic Acids Res*, **19**, 5619-5624.
- Palecek, E., Brazdova, M., Brazda, V., Palecek, J., Billova, S., Subramaniam, V. and Jovin, T.M. (2001) Binding of p53 and its core domain to supercoiled DNA. *Eur J Biochem*, **268**, 573-581.
- Palecek, E., Masarik, M., Kizek, R., Kuhlmeier, D., Hassmann, J. and Schulein, J. (2004) Sensitive electrochemical determination of unlabeled MutS protein and detection of point mutations in DNA. *Anal Chem*, **76**, 5930-5936.
- Pan, C., Xu, S., Zhou, H., Fu, Y., Ye, M. and Zou, H. (2007) Recent developments in methods and technology for analysis of biological samples by MALDI-TOF-MS. *Anal Bioanal Chem*, **387**, 193-204.
- Panigrahi, G.B., Lau, R., Montgomery, S.E., Leonard, M.R. and Pearson, C.E. (2005) Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. *Nat Struct Mol Biol*, **12**, 654-662.
- Paradkar, A., Trefzer, A., Chakraborty, R. and Stassi, D. (2003) Streptomyces genetics: a genomic perspective. *Crit Rev Biotechnol*, **23**, 1-27.
- Pearson, C.E., Ewel, A., Acharya, S., Fishel, R.A. and Sinden, R.R. (1997) Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum Mol Genet*, **6**, 1117-1123.
- Pearson, C.E. and Sinden, R.R. (1996) Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. *Biochemistry*, **35**, 5041-5053.
- Pearson, C.E., Tam, M., Wang, Y.H., Montgomery, S.E., Dar, A.C., Cleary, J.D. and Nichol, K. (2002) Slipped-strand DNAs formed by long (CAG)*(CTG) repeats: slipped-out repeats and slip-out junctions. *Nucleic Acids Res*, **30**, 4534-4547.
- Pitcher, R.S., Brissett, N.C. and Doherty, A.J. (2007) Nonhomologous end-joining in bacteria: a microbial perspective. *Annu Rev Microbiol*, **61**, 259-282.
- Pitcher, R.S., Wilson, T.E. and Doherty, A.J. (2005) New insights into NHEJ repair processes in prokaryotes. *Cell Cycle*, **4**, 675-678.
- Qin, Z. and Cohen, S.N. (2002) Survival mechanisms for Streptomyces linear replicons after telomere damage. *Mol Microbiol*, **45**, 785-794.
- Ramsden, D.A. and Gellert, M. (1998) Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *Embo J*, **17**, 609-614.
- Rass, U., Ahel, I. and West, S.C. (2007) Defective DNA repair and neurodegenerative disease. *Cell*, **130**, 991-1004.
- Redenbach, M., Kieser, H.M., Denapate, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A. (1996) A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome. *Mol Microbiol*, **21**, 77-96.
- Reed, M., Woelker, B., Wang, P., Wang, Y., Anderson, M.E. and Tegtmeyer, P. (1995) The C-terminal domain of p53 recognizes DNA damaged by ionizing radiation. *Proc Natl Acad Sci U S A*, **92**, 9455-9459.
- Resnick-Silverman, L., St Clair, S., Maurer, M., Zhao, K. and Manfredi, J.J. (1998) Identification of a novel class of genomic DNA-binding sites suggests a mechanism for selectivity in target gene activation by the tumor suppressor protein p53. *Genes Dev*, **12**, 2102-2107.

- Rhodes, D., Schwabe, J.W., Chapman, L. and Fairall, L. (1996) Towards an understanding of protein-DNA recognition. *Philos Trans R Soc Lond B Biol Sci*, **351**, 501-509.
- Riha, K., Heacock, M.L. and Shippen, D.E. (2006) The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology. *Annu Rev Genet*, **40**, 237-277.
- Riley, K.J., Ramirez-Alvarado, M. and Maher, L.J., 3rd. (2007) RNA-p53 interactions in vitro. *Biochemistry*, **46**, 2480-2487.
- Rocha, E.P., Cornet, E. and Michel, B. (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet*, **1**, e15.
- Romer, L., Klein, C., Dehner, A., Kessler, H. and Buchner, J. (2006) p53--a natural cancer killer: structural insights and therapeutic concepts. *Angew Chem Int Ed Engl*, **45**, 6440-6460.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning. A laboratory manual*. Spring Harbor Laboratory Press, New York.
- Sancar, A., Lindsey-Boltz, L.A., Unsal-Kacmaz, K. and Linn, S. (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem*, **73**, 39-85.
- SantaLucia, J., Jr. and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu Rev Biophys Biomol Struct*, **33**, 415-440.
- Savery, N.J. (2007) The molecular mechanism of transcription-coupled DNA repair. *Trends Microbiol*, **15**, 326-333.
- Sax, J.K. and El-Deiry, W.S. (2003) p53 downstream targets and chemosensitivity. *Cell Death Differ*, **10**, 413-417.
- Sayer, H.R. (2008) The DNA ligases and associated repair proteins of *Streptomyces coelicolor*. Thesis, PhD. University of East Anglia.
- Schar, P., Herrmann, G., Daly, G. and Lindahl, T. (1997) A newly identified DNA ligase of *Saccharomyces cerevisiae* involved in RAD52-independent repair of DNA double-strand breaks. *Genes Dev*, **11**, 1912-1924.
- Schildkraut, C. (1965) Dependence of the melting temperature of DNA on salt concentration. *Biopolymers*, **3**, 195-208.
- Seeman, N.C., Rosenberg, J.M. and Rich, A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc Natl Acad Sci U S A*, **73**, 804-808.
- Selivanova, G., Iotsova, V., Kiseleva, E., Strom, M., Bakalkin, G., Grafstrom, R.C. and Wiman, K.G. (1996) The single-stranded DNA end binding site of p53 coincides with the C-terminal regulatory region. *Nucleic Acids Res*, **24**, 3560-3567.
- Sengupta, S. and Harris, C.C. (2005) p53: traffic cop at the crossroads of DNA repair and recombination. *Nat Rev Mol Cell Biol*, **6**, 44-55.
- Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Boucherie, H. and Mann, M. (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci U S A*, **93**, 14440-14445.
- Shu, K.-X., Li, B. and Wu, L.-X. (2006) The p53 network: p53 and its downstream genes. *Colloids and Surfaces*.
- Shuck, S.C., Short, E.A. and Turchi, J.J. (2008) Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. *Cell Res*, **18**, 64-72.
- Shuman, S. (1994) Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *Journal of Biological Chemistry*, **269**, 32678-32684.
- Somasundaram, K. (2000) Tumor suppressor p53: regulation and function. *Front Biosci*, **5**, D424-437.
- Stansel, R.M., Subramanian, D. and Griffith, J.D. (2002) p53 binds telomeric single strand overhangs and t-loop junctions in vitro. *J Biol Chem*, **277**, 11625-11628.
- Stark, W.M., Boocock, M.R. and Sherratt, D.J. (1992) Catalysis by site-specific recombinases. *Trends Genet*, **8**, 432-439.

- Stenger, J.E., Tegtmeier, P., Mayr, G.A., Reed, M., Wang, Y., Wang, P., Hough, P.V. and Mastrangelo, I.A. (1994) p53 oligomerization and DNA looping are linked with transcriptional activation. *Embo J*, **13**, 6011-6020.
- Stros, M., Muselikova-Polanska, E., Pospisilova, S. and Strauss, F. (2004) High-affinity binding of tumor-suppressor protein p53 and HMGB1 to hemicatenated DNA loops. *Biochemistry*, **43**, 7215-7225.
- Subramanian, D. and Griffith, J.D. (2002) Interactions between p53, hMSH2-hMSH6 and HMG I(Y) on Holliday junctions and bulged bases. *Nucleic Acids Res*, **30**, 2427-2434.
- Subramanian, D. and Griffith, J.D. (2005) Modulation of p53 binding to Holliday junctions and 3-cytosine bulges by phosphorylation events. *Biochemistry*, **44**, 2536-2544.
- Tait, R.C., Rodriguez, R.L. and West, R.W., Jr. (1980) The rapid purification of T4 DNA ligase from a lambda T4 lig lysogen. *J Biol Chem*, **255**, 813-815.
- Teilum, K., Olson, J.G. and Kragelund, B.B. (2009) Functional aspects of protein flexibility. *Cell Mol Life Sci*, **66**, 2231-2247.
- Tyers, M. and Buttner, M.J. (2004) Growth and development the whole picture begins to emerge. *Curr Opin Microbiol*, **7**, 561-564.
- Umar, A., Boyer, J.C. and Kunkel, T.A. (1994) DNA loop repair by human cell extracts. *Science*, **266**, 814-816.
- Usdin, K. and Woodford, K.J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Res*, **23**, 4202-4209.
- Vogelstein, B., Lane, D. and Levine, A.J. (2000) Surfing the p53 network. *Nature*, **408**, 307-310.
- Volff, J.N. and Altenbuchner, J. (1998) Genetic instability of the Streptomyces chromosome. *Mol Microbiol*, **27**, 239-246.
- von Hippel, P.H. (1994) Protein-DNA recognition: new perspectives and underlying themes. *Science*, **263**, 769-770.
- von Hippel, P.H. (2007) From "simple" DNA-protein interactions to the macromolecular machines of gene expression. *Annu Rev Biophys Biomol Struct*, **36**, 79-105.
- Vousden, K.H. (2006) Outcomes of p53 activation--spoiled for choice. *J Cell Sci*, **119**, 5015-5020.
- Vousden, K.H. and Lane, D.P. (2007) p53 in health and disease. *Nat Rev Mol Cell Biol*, **8**, 275-283.
- Walker, J.R., Corpina, R.A. and Goldberg, J. (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, **412**, 607-614.
- Walter, K., Warnecke, G., Bowater, R., Deppert, W. and Kim, E. (2005) tumor suppressor p53 binds with high affinity to CTG.CAG trinucleotide repeats and induces topological alterations in mismatched duplexes. *J Biol Chem*, **280**, 42497-42507.
- Wang, B., Xiao, Z. and Ren, E.C. (2009a) Redefining the p53 response element. *Proc Natl Acad Sci U S A*, **106**, 14373-14378.
- Wang, G. and Vasquez, K.M. (2004) Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proc Natl Acad Sci U S A*, **101**, 13448-13453.
- Wang, G. and Vasquez, K.M. (2009) Models for chromosomal replication-independent non-B DNA structure-induced genetic instability. *Mol Carcinog*, **48**, 286-298.
- Wang, S.T., Setlow, B., Conlon, E.M., Lyon, J.L., Imamura, D., Sato, T., Setlow, P., Losick, R. and Eichenberger, P. (2006) The forespore line of gene expression in *Bacillus subtilis*. *J Mol Biol*, **358**, 16-37.
- Wang, Y., Ghosh, G. and Hendrickson, E.A. (2009b) Ku86 represses lethal telomere deletion events in human somatic cells. *Proc Natl Acad Sci U S A*, **106**, 12430-12435.
- Watson, J., Hays, F.A. and Ho, P.S. (2004) Definitions and analysis of DNA Holliday junction geometry. *Nucleic Acids Res*, **32**, 3017-3027.
- Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, **171**, 737-738.

- Weinberg, R.L., Veprintsev, D.B., Bycroft, M. and Fersht, A.R. (2005) Comparative binding of p53 to its promoter and DNA recognition elements. *J Mol Biol*, **348**, 589-596.
- Weller, G.R. and Doherty, A.J. (2001) A family of DNA repair ligases in bacteria? *FEBS Lett*, **505**, 340-342.
- Weller, G.R., Kysela, B., Roy, R., Tonkin, L.M., Scanlan, E., Della, M., Devine, S.K., Day, J.P., Wilkinson, A., d'Adda di Fagagna, F., Devine, K.M., Bowater, R.P., Jeggo, P.A., Jackson, S.P. and Doherty, A.J. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science*, **297**, 1686-1689.
- Wells, R.D., Dere, R., Hebert, M.L., Napierala, M. and Son, L.S. (2005) Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic Acids Res*, **33**, 3785-3798.
- Weterings, E. and Chen, D.J. (2008) The endless tale of non-homologous end-joining. *Cell Res*, **18**, 114-124.
- Wieland, G., Hemmerich, P., Koch, M., Stoyan, T., Hegemann, J. and Diekmann, S. (2001) Determination of the binding constants of the centromere protein Cbf1 to all 16 centromere DNAs of *Saccharomyces cerevisiae*. *Nucleic Acids Res*, **29**, 1054-1060.
- Wiesmuller, L. (2001) Genetic Stabilization by p53 Involves Growth Regulatory and Repair Pathways. *J Biomed Biotechnol*, **1**, 7-10.
- Wilkinson, A., Day, J. and Bowater, R. (2001) Bacterial DNA ligases. *Mol Microbiol*, **40**, 1241-1248.
- Willis, A., Jung, E.J., Wakefield, T. and Chen, X. (2004) Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. *Oncogene*, **23**, 2330-2338.
- Wilson, T.E., Topper, L.M. and Palmbo, P.L. (2003) Non-homologous end-joining: bacteria join the chromosome breakdance. *Trends Biochem Sci*, **28**, 62-66.
- Winter, R.B., Berg, O.G. and von Hippel, P.H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *Escherichia coli* lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry*, **20**, 6961-6977.
- Wohrle, D., Kennerknecht, I., Wolf, M., Enders, H., Schwemmle, S. and Steinbach, P. (1995) Heterogeneity of DM kinase repeat expansion in different fetal tissues and further expansion during cell proliferation in vitro: evidence for a casual involvement of methyl-directed DNA mismatch repair in triplet repeat stability. *Hum Mol Genet*, **4**, 1147-1153.
- Wojciechowska, M., Bacolla, A., Larson, J.E. and Wells, R.D. (2005) The myotonic dystrophy type 1 triplet repeat sequence induces gross deletions and inversions. *J Biol Chem*, **280**, 941-952.
- Wolcke, J., Reimann, M., Klumpp, M., Gohler, T., Kim, E. and Deppert, W. (2003) Analysis of p53 "latency" and "activation" by fluorescence correlation spectroscopy. Evidence for different modes of high affinity DNA binding. *J Biol Chem*, **278**, 32587-32595.
- Woods, D.B. and Vousden, K.H. (2001) Regulation of p53 function. *Exp Cell Res*, **264**, 56-66.
- Xu, J. and Morris, G.F. (1999) p53-mediated regulation of proliferating cell nuclear antigen expression in cells exposed to ionizing radiation. *Mol Cell Biol*, **19**, 12-20.
- Yakovleva, T., Pramanik, A., Kawasaki, T., Tan-No, K., Gileva, I., Lindegren, H., Langel, U., Ekstrom, T.J., Rigler, R., Terenius, L. and Bakalkin, G. (2001) p53 Latency. C-terminal domain prevents binding of p53 core to target but not to nonspecific DNA sequences. *J Biol Chem*, **276**, 15650-15658.
- Yang, C.C., Huang, C.H., Li, C.Y., Tsay, Y.G., Lee, S.C. and Chen, C.W. (2002) The terminal proteins of linear *Streptomyces* chromosomes and plasmids: a novel class of replication priming proteins. *Mol Microbiol*, **43**, 297-305.
- Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103-119.

- Yeh, C.S., Chen, F.M., Wang, J.Y., Cheng, T.L., Hwang, M.J. and Tzou, W.S. (2003) Directional shape complementarity at the protein-DNA interface. *J Mol Recognit*, **16**, 213-222.
- Yoo, S. and Dynan, W.S. (1998) Characterization of the RNA binding properties of Ku protein. *Biochemistry*, **37**, 1336-1343.
- Yu, J. and Zhang, L. (2005) The transcriptional targets of p53 in apoptosis control. *Biochem Biophys Res Commun*, **331**, 851-858.
- Zahrt, T.C. and Deretic, V. (2000) An essential two-component signal transduction system in Mycobacterium tuberculosis. *J Bacteriol*, **182**, 3832-3838.
- Zhang, N., Gorin, A., Majumdar, A., Kettani, A., Chernichenko, N., Skripkin, E. and Patel, D.J. (2001) Dimeric DNA quadruplex containing major groove-aligned A-T-A-T and G-C-G-C tetrads stabilized by inter-subunit Watson-Crick A-T and G-C pairs. *J Mol Biol*, **312**, 1073-1088.
- Zhang, Z., Hu, W., Cano, L., Lee, T.D., Chen, D.J. and Chen, Y. (2004) Solution structure of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions. *Structure*, **12**, 495-502.
- Zhu, H. and Shuman, S. (2009) Gap filling activities of Pseudomonas LigD polymerase and functional interactions of LigD with the DNA end-binding Ku protein. *J Biol Chem*.
- Zierhut, C. and Diffley, J.F. (2008) Break dosage, cell cycle stage and DNA replication influence DNA double strand break response. *Embo J*, **27**, 1875-1885.
- Zlatanova, J. and van Holde, K. (1998) Binding to four-way junction DNA: a common property of architectural proteins? *Faseb J*, **12**, 421-431.

Appendix I

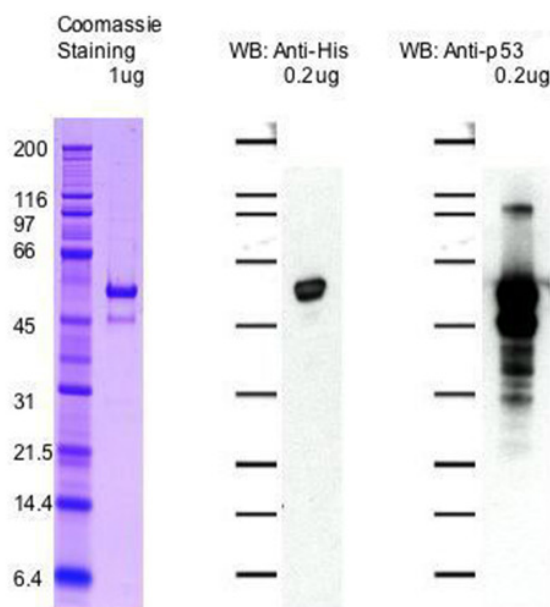


Figure AI.1 Recombinant human p53 analysis (Supplied by Active Motif). Human p53 (referred to in text as p53-B) was over-expressed in *E. coli* and purified via affinity chromatography. Following SDS-PAGE, purified protein was analysed by coomassie staining and western blot (directed against an NTD polyhistidine tag (Anti-His) and against the CTD (Anti-p53)).

Appendix II

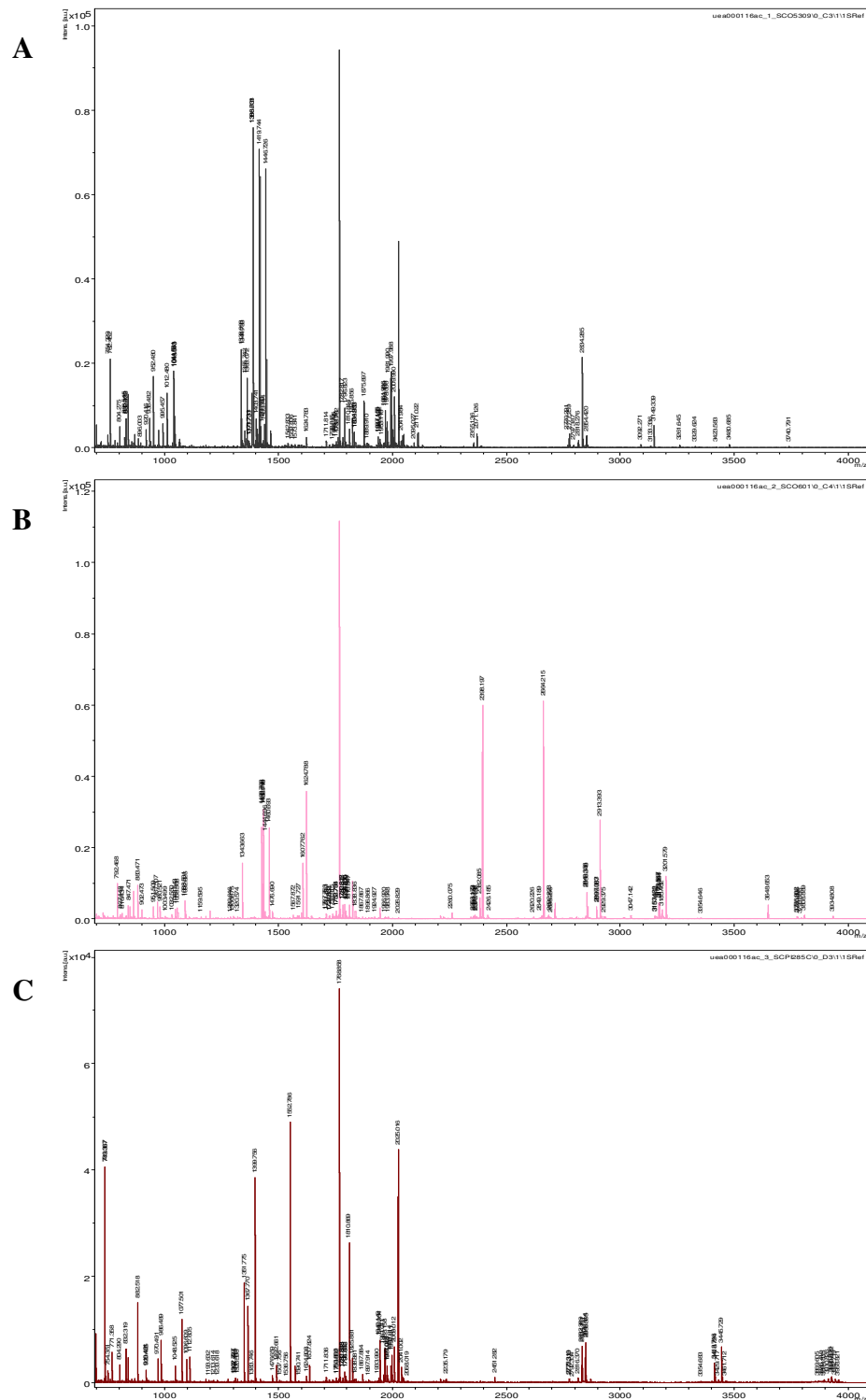


Figure AII.1 ScKu protein data collected from MALD-TOF spectrometry. For confirmation of identities of proteins over-expressed and purified during this study, mass spectrometry analysis was performed that verified the recombinant proteins were as expected. **A.** ScKu1 peak data. **B.** ScKu2 peak data. **C.** ScKu3 peak data.

Appendix III

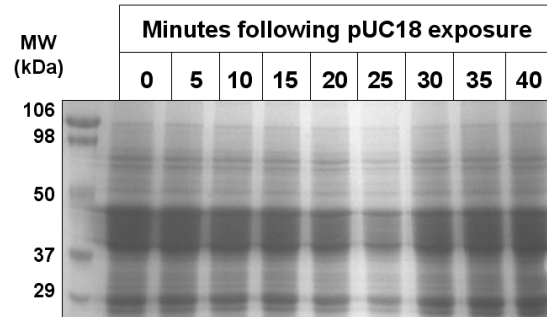


Figure AIII.1 SDS-PAGE and coomassie staining analysis of MtrA degradation following pUC18 exposure. 1 μ l of 100 ng / μ l of plasmid sample was added to 35 μ g of MtrA in a final volume of 30 μ l and incubated for up to 40 minutes.

Appendix IV

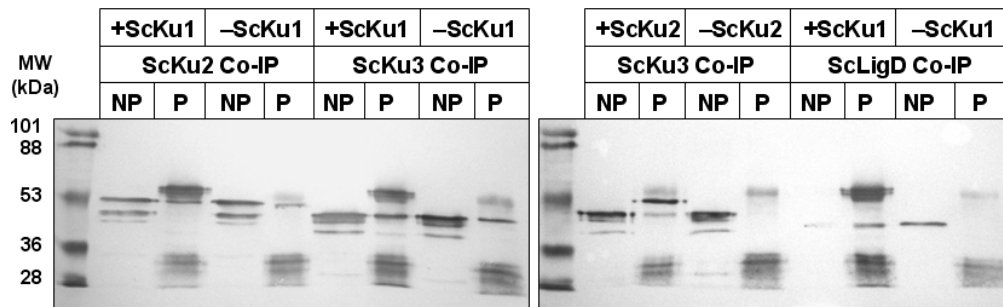


Figure AIV.1 Example of SDS-PAGE and western blot analysis of *S. coelicolor* NHEJ components following Co-IP protein-protein interaction analysis. Assays were performed using antibodies specific for ScKu1 and ScKu2 to immunoprecipitate ScKu3 or ScLigD. As controls, assays were performed without ScKu1 and ScKu2 to determine non-specific precipitation by antibodies. NP – Non-precipitated protein, P – Precipitated protein.

Appendix V

Table AV.1 *In silico* analysis of Ku homology searches in *Streptomyces* organisms that have been sequenced. Amino acid sequences for ScKu1, ScKu2 and ScKu3 were taken from StrepDB website and 'BLASTP' was used to identify homologues in *Streptomyces avermitilis*, *Streptomyces griseus* and *Streptomyces scabies*. Sequence identities above 40% are shown.

<i>S. coelicolor</i> <i>ku</i> gene	<i>S. avermitilis</i>	<i>S. griseus</i>	<i>S. scabies</i>
<i>scku1</i>	SAV_2945 ku1 (71%) SAV_879 ku2 (53%)	SGR_2195 (64%) SGR_6689 (49%)	SCAB29491 (67%)
<i>scku2</i>	-	-	-
<i>scku3</i>	SAV_879 ku2 (70%) SAV_2945 ku1 (49%)	SGR_6689 (60%) SGR_2195 (49%)	SCAB29491 (51%)